N 69 10611
NASA CR 97485

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-52, Vol. III

Supporting Research and Advanced Development

For the Period June 1 to July 31, 1968

CASE FILE
COPY

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

August 31, 1968

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-52, Vol. III

Supporting Research and Advanced Development

For the Period June 1 to July 31, 1968

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

August 31, 1968

SPACE PROGRAMS SUMMARY 37-52, VOL. III

Copyright © 1968
Jet Propulsion Laboratory
California Institute of Technology

Prepared Under Contract No. NAS 7-100
National Aeronautics & Space Administration

Preface

The Space Programs Summary is a bimonthly publication that presents a review of engineering and scientific work performed, or managed, by the Jet Propulsion Laboratory for the National Aeronautics and Space Administration during a two-month period. Beginning with the 37-47 series, the Space Programs Summary is composed of four volumes:

Vol. I. *Flight Projects* (Unclassified)

Vol. II. *The Deep Space Network* (Unclassified)

Vol. III. *Supporting Research and Advanced Development* (Unclassified)

Vol. IV. *Flight Projects and Supporting Research and Advanced Development* (Confidential)

Approved by:

W. H. Pickering, Director
Jet Propulsion Laboratory

Contents

SYSTEMS DIVISION

I. Systems Analysis Research 1

A. Effect on Precession Parameters of New Planetary Masses Used in Ephemeris Development
J. D. Mulholland, NASA Code 129-04-04-02 1

B. On the Partial Derivatives of the Two-Body Problem
R. A. Broucke, NASA Code 129-04-01-02 2

II. Computation and Analysis 10

A. LEASTQ: A Program for Least-Squares Fitting Segmented Cubic Polynomials Having a Continuous First Derivative
T. M. Lang, NASA Code 129-04-04-01 10

B. Abstracts of Certain Mathematical Subroutines, III. Minimization Subject to Linear Inequality and Equality Constraints With an Application to Curve Fitting
R. J. Hanson and A. J. Semtner, NASA Code129-04-04-01 15

PROJECT ENGINEERING DIVISION

III. System Design and Integration 21

A. Entry and Landing Capsule System
E. K. Casani, NASA Code 186-68-09-09 21

GUIDANCE AND CONTROL DIVISION

IV. Flight Computers and Sequencers 27

A. STAR Computer Assembler and Loader
J. A. Rohr, NASA Code 125-17-04-03 27

V. Spacecraft Power 31

A. Thermionic Converter Technology
O. S. Merrill, NASA Code 120-33-02-06 31

B. Low Saturation Drop Transistor
A. I. Schloss, NASA Code 120-33-08-05 37

C. A Cell for the Direct Observation of Gassing Phenomena at Battery Electrode Surfaces in Low-Gravity Environments
G. L. Juvinall, NASA Code 120-34-01-07 38

D. The Products of the Electrochemical Oxidation of Zinc Battery Electrodes
G. L. Juvinall, NASA Code 120-34-01-17 44

E. Development of the Heat Sterilizable Battery
R. Lutwack, NASA Code 120-34-01-03, -05, -06, -10, -12, -13, -14, -18 45

Contents (contd)

F. RTG Test Laboratory
R. G. Ivanoff, NASA Code 120-27-06-01 46

G. Analog Voltage to Duty Cycle Generator
A. I. Schloss, NASA Code 120-33-08-03 47

VI. Spacecraft Control 49

A. Development of an Approach-Guidance Optical Planet Tracker
F. R. Chamberlain, NASA Code 186-68-02-23 47

B. Strapdown Inertial System Analysis
G. Paine, NASA Code 125-17-01-04 52

C. Strapdown Inertial System Alignment Technique
B. R. Markiewicz, NASA Code 125-17-01-04. 55

D. Strapdown Electrostatically-Suspended Gyro Drift Math Model Development
V. A. Karpenko and D. H. Lipscomb, NASA Code 125-17-01-04 58

VII. Guidance and Control Research 65

A. Emitter Work Function of an Operational Converter
K. Shimada, NASA Code 129-02-01-07 65

B. Surface Barriers on Layer Semiconductors: GaSe
S. Kurtin and C. A. Mead, NASA Code 129-02-05-09 68

C. Noise Measurements on a Double-Injection Silicon Diode
D. H. Lee, H. R. Bilger, and M-A. Nicolet, NASA Code 129-02-05-09 70

ENGINEERING MECHANICS DIVISION

VIII. Electronic Packaging and Cabling 75

A. Thermal Resistance of Transistors in JEDEC TO-5 and TO-18 Packages
R. M. Jorgensen, NASA Code 125-25-03-01 75

B. Simplifying Complex Miniature Interconnections
L. Katzin, NASA Code 125-25-03-02 80

C. Documentation of Wiring Harnesses Using Punched-Card Techniques
W. G. Kloezeman, NASA Code 125-25-03-01 84

PROPULSION DIVISION

IX. Solid Propellant Engineering 87

A. Study of Radiative Cooling of a Graphite Composite Nozzle
S. Fogler, NASA Code 180-32-07-01 87

B. Solid Propellant Rocket Motor Command Termination by Water Injection
L. D. Strand, NASA Code 731-26-02-02. 89

Contents (contd)

X. Polymer Research 97

A. Saturated Hydrocarbon Prepolymers
J. D. Ingham, NASA Code 128-32-05-03 97

B. Investigation on Sterilizable Polymer Battery Separators, Part II
E. F. Cuddihy and J. Moacanin, NASA Code 120-34-01-21 98

C. The Ethylene Oxide–Freon 12 Decontamination Procedure: Control and Determination of the Moisture Content of the Chamber
R. H. Silver and S. H. Kalfayan, NASA Code 186-58-13-09 101

XI. Research and Advanced Concepts 106

A. Special Applications for Spectroscopic Scanning of Internal Plasma Flows
E. J. Roschke, NASA Code 129-01-05-10 106

B. Hall and Ion-Slip Effects in Channel Flow
E. J. Roschke, NASA Code 129-01-05-11 109

C. Liquid-Metal MHD Power Conversion
D. G. Elliott and L. G. Hays, NASA Code 120-27-06-03 113

D. Dynamic Gas Effects on the Breakdown Potential of Helium
J. A. Gardner, NASA Code 129-01-05-11 119

E. Lithium–Boiling Potassium Test Loop Runs With Reactor Simulator
H. Gronroos and G. Kikin, NASA Code 120-27-06-14 121

XII. Liquid Propulsion 128

A. The Reaction of OF_2 With B_2H_6: Rate of Formation of BF_3 and Consumption Rates of OF_2 and B_2H_6
R. A. Rhein, NASA Code 128-31-52-01 128

SPACE SCIENCES DIVISION

XIII. Space Instruments 135

A. High- and Low-Field Operation of the Helium Vector Magnetometer
F. E. Vescelus, NASA Code 188-36-01-04 135

B. Ground Support System for X-Ray/Gamma-Ray Prototype Flight Experiment
L. L. Lewyn, NASA Code 188-41-01-01 138

C. Parallel to Serial Converter for Punch Tape Perforation
L. L. Lewyn, NASA Code 188-41-01-01 139

D. Electronics System for Measuring Induced Photon Spectra at the UCLRL Bevatron
L. L. Lewyn, NASA Code 185-42-13-01 139

Contents (contd)

E. Developmental Flight-Model Infrared Interferometer
R. A. Schindler, NASA Code 185-37-32-01 142

XIV. Science Data Systems 149

A. Decomposition of the States of a Linear Feedback Shift Register Into Cycles of Equal Length
M. Perlman, NASA Code 125-23-02-02 149

B. Capsule System Advanced Development Entry Data Subsystem
R. V. Gutierrez, NASA Code 186-68-03-04 154

XV. Lunar and Planetary Instruments and Sciences 162

A. Infrared Absorption Spectrum of CH_4 at 9050 cm^{-1}
J. S. Margolis and K. Fox, NASA Code 185-41-34-01 162

B. Peak-Detector–Analog/Pulse Width Converter Analysis
J. R. Locke, NASA Code 125-24-01-08 163

XVI. Bioscience 172

A. Soil Studies—Desert Microflora. XV. Analysis of Antarctic Dry Valley Soils by Cultural and Radiorespirometric Methods
J. S. Hubbard, R. E. Cameron, and A. B. Miller, NASA Code 189-55-04-03 172

XVII. Fluid Physics 176

A. Magnetic Topology and Flux Reconnection
A. Bratenahl and C. Yeates, NASA Code 129-02-08-04 176

XVIII. Physics 183

A. Enhanced Fluctuations in a Plasma Due to Ion Cyclotron Instability
C.-S. Wu and J. S. Zmuidzinas, NASA Code 129-02-07-02 183

TELECOMMUNICATIONS DIVISION

XIX. Communications Systems Research 185

A. Coding and Synchronization Studies: Effect of Quantization on the *Mariner* Mars 1969 Digital Dump Matched Filter
J. K. Holmes, NASA Code 125-21-02-03 185

B. Combinatorial Communication: Orthogonal Codes and Erasures
L. R. Welch, NASA Code 125-21-01-01 187

C. Combinatorial Communication: Sphere-Packing in the Hamming Metric
R. J. McEliece and H. Rumsey, Jr., NASA Code 125-21-01-01 189

Contents (contd)

D. Combinatorial Communication: A General Formulation of Error Metrics
S. W. Golomb, NASA Code 125-21-01-01 190

E. Combinatorial Communication: The Maximum Number of Cycles in the de Bruijn Graph
H. Fredricksen, NASA Code 125-21-01-01 193

F. Information Processing: Prediction With Piecewise Linear Correlation Functions
I. F. Blake, NASA Code 150-22-11-09 197

G. Information Processing: Least-Squares Estimates From Likelihood Ratios
T. Kailath, NASA Code 150-22-11-09 199

H. Astrometrics: Toeplitz Matrix Inversion—The Algorithm of W. F. Trench
S. Zohar, NASA Code 150-22-11-10 203

I. Data Compression Techniques: Entropy of Graphs
E. C. Posner, NASA Code 150-22-17-08. 209

XX. Communications Elements Research 216

A. Spacecraft Antenna Research: Radiation From a Turnstile Antenna Located in a Plasma Shell
R. Woo, NASA Code 125-22-01-02 216

B. Spacecraft Antenna Research: Antenna Tolerances
R. W. Dickinson, NASA Code 186-68-04-02 219

C. Precision Calibration Techniques: Microwave Thermal Noise Standards
C. T. Stelzried, NASA Code 150-22-11-07 223

D. A Precision DC Potentiometer Insertion Loss Test Set and Reflectometer for Use at 90 GHz
D. A. Oltmans and T. Sato, NASA Code 125-21-03-04 229

E. Accuracy of Numerically Computed Electromagnetic Scattered Patterns
S. A. Brunstein, R. E. Cormack, and A. C. Ludwig, NASA Code 150-22-13-12 . . . 233

XXI. Spacecraft Telemetry and Command 239

A. Frequency Acquisition in an MFSK Receiver
H. Chadwick, NASA Code 150-22-17-04 239

XXII. Spacecraft Radio. 249

A. Low Data Rate Telemetry RF System Development
R. Postal, NASA Code 150-22-17-06 249

B. RF Power Amplifier Life Test Summary
R. Hughes, NASA Code 186-68-04-09 250

Contents (contd)

ADVANCED STUDIES

XXIII. Future Projects 253

A. The Objectives for Roving Vehicles in a Lunar Exploration Program

R. G. Brereton, NASA Code 945-41-00-00 253

I. Systems Analysis Research

SYSTEMS DIVISION

A. Effect on Precession Parameters of New Planetary Masses Used In Ephemeris Development, *J. D. Mulholland*

In a previous article (SPS 37-45, Vol. IV, pp. 17–19), a new set of planetary masses was proposed for use in ephemeris development computations. This set, with some slight modifications, has since been incorporated into a formally recommended set of JPL astrodynamic constants (Ref. 1) and is now in use in the ephemeris development effort. It is believed that these represent a substantially more reliable set of values than the IAU values, which date from the time of Simon Newcomb. Still, the interplay between the various quantities that are used to describe the physical universe is frequently indirect enough that one must be cautious about introducing unwanted problems in one area by the very act of improving the accuracy somewhere else in the system.

The perturbative effect of the planets on the spatial orientation of the ecliptic appears in the planetary precession. Thus, if the planetary masses are changed, this should imply corresponding modifications to the precession parameters. Indeed, Lieske (Ref. 2) has derived and published the partial derivatives of the precession parameters with respect to the planetary masses.

On the other hand, the appropriate modifications have not been applied to the precession parameters in Ref. 1 nor in any of our programs. Before this state of affairs is explained, it is best to examine the values that such corrections would take. Table 1 lists the results of evaluating Lieske's expressions for the mass system in Ref. 1. The

Table 1. Corrections to the precession parameters

Parameter	Coefficient of				
	1	T_1	T	$T_1 T$	T^2
General formulae					
$\Delta\bar{\epsilon}_0$	—	+0.″038	—	—	—
$\Delta\bar{\epsilon}$	—	+0.038	+0.038	—	—
ΔP	—	—	—	−0.002	−0.001
$\Delta\zeta_0$	—	—	+0.002	−0.001	—
ΔZ	—	—	+0.002	−0.001	−0.001
$\Delta\theta$	—	—	−0.010	+0.001	—
Evaluated for T_1 = 1950.0					
$\Delta\bar{\epsilon}_0$	+0.019	—	—	—	—
$\Delta\bar{\epsilon}$	+0.019	—	+0.038	—	—
ΔP	—	—	−0.001	—	−0.001
$\Delta\zeta_0$	—	—	+0.002	—	—
ΔZ	—	—	+0.002	—	−0.001
$\Delta\theta$	—	—	−0.010	—	—

values given in the lower part of Table 1 are those that correspond to the expressions for ζ_0, Z, θ, $\bar{\epsilon}$ given in Ref. 1.

It is now known (cf, e.g., Ref. 3) that the conventional value for the constant of general precession is too small by some amount in the range $0''.6 \leq \Delta P \leq 1''.3$, and it appears that the uncertainty in the value of the mean obliquity (Ref. 4) is in the vicinity of $-0''.3 \leq (\Delta\bar{\epsilon} + 0.3T) \leq -0''.1$ also. If we hypothesize that the "true" errors are $1''.0$ and $-0''.2$, respectively, then this implies the following errors in the precession parameters:

$$\Delta\bar{\epsilon} = -0''.2$$

$$\Delta P = +1''.0$$

$$\Delta\zeta_0 = +0.460\,T$$

$$\Delta Z = +0.460\,T$$

$$\Delta\theta = +0.402\,T$$

These numbers are much larger than those in the lower half of Table 1, which means that the inconsistency involved in retaining the conventional values for the precession parameters is notably smaller than the known error that contaminates the entire system. This is not taken as a warrant to ignore the existence of the inconsistency and its possible consequences. Indeed, the entire subject of precession is now under study. These results are, however, interpreted as a justification of the adoption of mass values that better satisfy the observed motions in the solar system at the earliest date, even at the cost of what is hoped to be a temporary inconsistency.

References

1. Melbourne, W. G., Mulholland, J. D., Sjogren, W. L., and Sturms, F. M., *Constants and Related Information for Astrodynamic Calculations, 1968*, Technical Report 32-1306. Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1968.

2. Lieske, J., *Expressions for the Precession Quantities and Their Partial Derivatives*, Technical Report 32-1044. Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1967.

3. Fricke, W., "Precession and Galactic Rotation Derived from Fundamental Proper Motions of Distant Stars," *Astron. J.*, Vol. 72, pp. 1368–1379, Dec. 1967.

4. Duncombe, R. L., "Motion of Venus 1750–1949," *Astron. Papers of the Am. Ephemeris*, Vol. XVI, Part 1, U.S. Government Printing Office, Washington, D. C., 1958.

B. On the Partial Derivatives of the Two-Body Problem, *R. A. Broucke*

1. Introduction

It is well known that in the differential correction of orbits of celestial bodies, with the use of observations, the partial derivatives of the two-body problem play an important role. It is also known that these partial derivatives are of extreme importance in the computation of perturbations on the celestial objects due to other massive bodies. This has been recently shown in detail by this author (SPS 37-49, Vol. III, pp. 31–40).

The problem which is crucial in the computation of perturbations, i.e., the development of planetary theories, is to find good explicit forms for Green's functions or the fundamental matrices that are required. This is shown in a fundamental article published by J. M. A. Danby in 1962 (Ref. 1), which studies the foundations of the problems associated with generating planetary theories. Our article attempts only to develop in detail some technical formulas necessary for an implementation of Danby's theory on the computer. We have work in progress to implement this theory.

Since the three-dimensional Kepler problem is represented by a sixth-order system of differential equations, the solution will depend on six arbitrary constants. We are interested in the matrix with the $6 \times 6 = 36$ partials of the position in phase space with respect to the six integration constants, and also the inverse of this matrix. For the six variables in phase space, we will use the three position coordinates x, y, z and three velocity components $\dot{x}, \dot{y}, \dot{z}$ with respect to some given inertial frame of reference.

We will develop the results for three different sets of the six integration constants. The first set will be the classical system of orbit elements $(a, e, M_0, i, \omega, \Omega)$. The others will be the set-I and set-III elements defined by Brouwer and Clemence (Ref. 2, pp. 238 and 241). Our results are derived in such a way that the partial derivatives with the set-III elements are obtained from those with the classical elements by simple matrix operations. We have obtained our results by rather elementary and straightforward algebraic development, mainly because we intend to extend our results later to some other sets of variables and orbit elements, by applying a chain of successive simple changes of variables.

The principal result presented in this article is a consistent derivation of four fundamental solution matrices

(Green's functions) of the variational equations of the two-body problem. The fundamental matrices are Matrices (10), (25), (31), and (37). In Expression (13), we indicate the basic formula to obtain the inverse of the fundamental matrix. As an application, the inverse of Matrix (10) is given explicitly in Expression (14).

2. The Basic Formulation of Kepler's Problem

We will use the formulation of the two-body solution which makes use of the three classical orthonormal vectors $(\mathbf{P}, \mathbf{Q}, \mathbf{R})$ depending on the three orbit elements (i, ω, Ω) only. These vectors are well known (Ref. 2, p. 33) so their components are not reproduced here.

The position and velocity vectors $\mathbf{x}$ and $\dot{\mathbf{x}}$ depend on (i, ω, Ω) through the components of $\mathbf{P}$ and $\mathbf{Q}$, and on the other three orbit elements through the intermediate quantities X, Y:

$$X = a(\cos E - e), \qquad Y = a(1 - e^2)^{1/2} \sin E \tag{1}$$

The expressions for the position and velocity vectors are

$$\mathbf{x} = \mathbf{P}X + \mathbf{Q}Y, \qquad \dot{\mathbf{x}} = \mathbf{P}\dot{X} + \mathbf{Q}\dot{Y} \tag{2}$$

The time derivatives of X and Y are

$$\dot{X} = \frac{-na^2 \sin E}{r}, \qquad \dot{Y} = \frac{na^2(1 - e^2)^{1/2} \cos E}{r} \tag{3}$$

The notations used in the above formulas are classical, and the meanings of most of our symbols are known. We have used E for the eccentric anomaly. We will also use the symbol $\bar{M}$ for the mean anomaly and r for the radius:

$$\bar{M} = nt + M_0 = E - e \sin E, \qquad r = a(1 - e \cos E) \tag{4}$$

The time derivatives of $\bar{M}$, E, and r are given by

$$\left.\begin{aligned} \dot{\bar{M}} &= n \\ \dot{E} &= \frac{na}{r} \\ \dot{r} &= \frac{a^2 e n \sin E}{r} \end{aligned}\right\} \tag{5}$$

The mean motion n and the semimajor axis a are related by

$$n^2 a^3 = \mu, \qquad \frac{\partial a}{\partial n} = -\frac{2a}{3n} \tag{6}$$

The partial derivatives of some important quantities are given in Expression (7). The derivation of these quantities is straightforward. In deriving the partials of the two-body equations, Expression (7), we consider them as being a set of functions depending on the time t (measured from a given epoch) and six parameters. The six

Expression (7)

	$/\partial a$	$/\partial e$	$/\partial M_0$
$\partial r/$	$\frac{r}{a} - \frac{3aent \sin E}{2r}$	$-a \cos E + \frac{a^2 e \sin^2 E}{r}$	$\frac{a^2 e \sin E}{r}$
$\partial \bar{M}/$	$\frac{-3nt}{2a}$	0	$+1$
$\partial E/$	$\frac{-3nt}{2r}$	$\frac{a \sin E}{r}$	$\frac{a}{r}$
$\partial X/$	$\frac{1}{a}\left(X - \frac{3t}{2}\dot{X}\right)$	L	$\frac{\dot{X}}{n}$
$\partial Y/$	$\frac{1}{a}\left(Y - \frac{3t}{2}\dot{Y}\right)$	M	$\frac{\dot{Y}}{n}$
$\partial \dot{X}/$	$-\frac{1}{2a}\left(\dot{X} - 3\mu\frac{Xt}{r^3}\right)$	$\dot{L}$	$-n\left(\frac{a}{r}\right)^3 X$
$\partial \dot{Y}/$	$-\frac{1}{2a}\left(\dot{Y} - 3\mu\frac{Yt}{r^3}\right)$	$\dot{M}$	$-n\left(\frac{a}{r}\right)^3 Y$

parameters (or orbit elements) are considered independent. When the partial derivative with respect to one element is taken, the five other elements and the time are kept constants, but the eccentric anomaly E, mean anomaly $\bar{M}$, and radius r are treated as functions of t and of the orbit elements.

We have used the following auxiliary variables related to the partial derivatives with respect to the eccentricity e:

$$\left.\begin{aligned} L &= \frac{a^2}{r}(e\cos E - 1 - \sin^2 E) \\ M &= \frac{a^2 \sin E}{r(1-e^2)^{1/2}}(\cos E - e) \\ \dot{L} &= \frac{na^4}{r^3}(e - 2\cos E + e\cos^2 E)\sin E \\ \dot{M} &= \frac{na^4}{r^3(1-e^2)^{1/2}}(e^2 - 1 - e\cos E \\ &\quad + 2\cos^2 E - e\cos^3 E) \end{aligned}\right\} \tag{8}$$

All seven quantities on each line of Expression (7) are independent of the orbit elements i, ω, Ω, and the corresponding partial derivatives are thus zero. Only **P** and **Q** will give a contribution in taking the derivatives of **x** and $\dot{\mathbf{x}}$ with respect to (i, ω, Ω). For this reason, we will now introduce two 3×3 matrices with the nine partial derivatives of the components of **P** and **Q** with respect to the elements (i, ω, Ω). We will designate each matrix by its first element, written between brackets. The elements (i, ω, Ω) correspond to each column of the matrices, while the three rows refer to the three components of **P** or **Q**. The partials of **R** will not be needed.

We have

$$\left.\begin{aligned} \left[\frac{\partial \mathbf{P}}{\partial i}\right] &= \begin{bmatrix} +P_z \sin\Omega & Q_x & -P_y \\ -P_z\cos\Omega & Q_y & P_x \\ \sin\omega\cos i & Q_z & 0 \end{bmatrix} \\ \left[\frac{\partial \mathbf{Q}}{\partial i}\right] &= \begin{bmatrix} +Q_z\sin\Omega & -P_x & -Q_y \\ -Q_z\cos\Omega & -P_y & Q_x \\ \cos\omega\cos i & -P_z & 0 \end{bmatrix} \end{aligned}\right\} \tag{9}$$

It is obvious that in computing the derivatives of the position and velocity with respect to (a, e, M_0) all that is needed is Expression (7), while Expression (9) contains all the required information for the derivatives with respect to (i, ω, Ω). This separation of the six orbit elements in two groups of three greatly facilitates the computation of partial derivatives. Following is the 6×6 matrix of partials of **x** and $\dot{\mathbf{x}}$ with respect to the six orbit elements $(a, e, M_0, i, \omega, \Omega)$. The columns of the matrix correspond to these orbit elements in the order given, while the six rows are in the order $(x, y, z, \dot{x}, \dot{y}, \dot{z})$.

$$\left[\frac{\partial \mathbf{x}}{\partial a}\right] = \begin{array}{|c|c|c|l|l|l|} \hline a & e & M_0 & i & \omega & \Omega \\ \hline \multirow{3}{*}{$\frac{1}{a}\left(\mathbf{x} - \frac{3}{2}\dot{\mathbf{x}}t\right)$} & \multirow{3}{*}{$L\mathbf{P} + M\mathbf{Q}$} & \multirow{3}{*}{$\frac{\dot{\mathbf{x}}}{n}$} & $+z\sin\Omega$ & $\mathbf{Q}X - \mathbf{P}Y$ & $-y$ \\ & & & $-z\cos\Omega$ & $= \mathbf{R}\times\mathbf{x}$ & $+x$ \\ & & & $+(X\sin\omega + Y\cos\omega)\cos i$ & & 0 \\ \hline \multirow{3}{*}{$\frac{-1}{2a}\left(\dot{\mathbf{x}} - 3\mu\frac{\mathbf{x}t}{r^3}\right)$} & \multirow{3}{*}{$\dot{L}\mathbf{P} + \dot{M}\mathbf{Q}$} & \multirow{3}{*}{$-n\left(\frac{a}{r}\right)^3\mathbf{x}$} & $+\dot{z}\sin\Omega$ & $\mathbf{Q}\dot{X} - \mathbf{P}\dot{Y}$ & $-\dot{y}$ \\ & & & $-\dot{z}\cos\Omega$ & $= \mathbf{R}\times\dot{\mathbf{x}}$ & $+\dot{x}$ \\ & & & $+(\dot{X}\sin\omega + \dot{Y}\cos\omega)\cos i$ & & 0 \\ \hline \end{array} \tag{10}$$

3. The Partial Derivatives of the Six Orbit Elements

We will now develop the matrix with partial derivatives of the orbit elements with respect to the position and velocity. In other words, we will develop the inverse of Matrix (10). This inversion can be performed with the use of Lagrange brackets $[a_\alpha, a_\beta]$ and Poisson parentheses (a_α, a_β). These may be written in the following form:

$$\left.\begin{aligned}[[a_\alpha, a_\beta]] &= \left[\frac{\partial \mathbf{x}}{\partial a_\alpha}\right]^T S \left[\frac{\partial \mathbf{x}}{\partial a_\beta}\right] \\ [(a_\alpha, a_\beta)] &= \left[\frac{\partial a_\alpha}{\partial \mathbf{x}}\right] S \left[\frac{\partial a_\beta}{\partial \mathbf{x}}\right]^T\end{aligned}\right\} \tag{11}$$

where T indicates the transpose, and where S is a 6×6 matrix

$$S = \begin{vmatrix} 0 & +I_3 \\ -I_3 & 0 \end{vmatrix} \tag{12}$$

where I_3 is the 3×3 unit matrix.

The different Lagrange brackets and Poisson parentheses corresponding to the classical orbit elements are in many textbooks and we do not reproduce them here. As a consequence of the Definitions in Expression (11), we may directly write an expression for the inverse of the Matrix (10):

$$\left[\frac{\partial a_\alpha}{\partial \mathbf{x}}\right] = [(a_\alpha, a_\beta)]\left[\frac{\partial \mathbf{x}}{\partial a_\beta}\right]^T S^{-1} \tag{13}$$

The symbol a_α represents the six orbit elements used in *Subsection 2*. We have used the convention that Latin indices go from 1 to 3 and Greek indices from 1 to 6. Expression (14) gives the explicit terms for Expression (13).

Expression (14)

$\dfrac{\partial a}{\partial \mathbf{x}} = \dfrac{2a^2}{r^3}\,\mathbf{x}$	$\dfrac{\partial a}{\partial \dot{\mathbf{x}}} = \dfrac{2\dot{\mathbf{x}}}{n^2 a}$
$\dfrac{\partial e}{\partial \mathbf{x}} = \dfrac{(1-e^2)^{1/2}}{na^2 e}\left[\mathbf{Q}\dot{X} - \mathbf{P}\dot{Y} + n(1-e^2)^{1/2}\left(\dfrac{a}{r}\right)^3 \mathbf{x}\right]$	$\dfrac{\partial e}{\partial \dot{\mathbf{x}}} = \dfrac{(1-e^2)^{1/2}}{na^2 e}\left[\mathbf{P}Y - \mathbf{Q}X + (1-e^2)^{1/2}\dfrac{\dot{\mathbf{x}}}{n}\right]$
$\dfrac{\partial i}{\partial \mathbf{x}} = +\dfrac{\left[(\mathbf{P}\dot{Y} - \mathbf{Q}\dot{X})\cos i + \dfrac{\partial \dot{\mathbf{x}}}{\partial \Omega}\right]}{na^2(1-e^2)^{1/2}\sin i}$	$\dfrac{\partial i}{\partial \dot{\mathbf{x}}} = -\dfrac{\left[(\mathbf{P}Y - \mathbf{Q}X) + \dfrac{\partial \mathbf{x}}{\partial \Omega}\right]}{na^2(1-e^2)^{1/2}\sin i}$
$\dfrac{\partial M_0}{\partial \mathbf{x}} = \dfrac{+1}{na^2}\left[-\dot{\mathbf{x}} + 3\mu\dfrac{\mathbf{x}t}{r^3} + \dfrac{(1-e^2)}{e}(\dot{L}\mathbf{P} + \dot{M}\mathbf{Q})\right]$	$\dfrac{\partial M_0}{\partial \dot{\mathbf{x}}} = \dfrac{+1}{na^2}\left[-2\mathbf{x} + 3\dot{\mathbf{x}}t - \dfrac{1-e^2}{e}(L\mathbf{P} + M\mathbf{Q})\right]$
$\dfrac{\partial \omega}{\partial \mathbf{x}} = \dfrac{1}{na^2}\left[\dfrac{-(1-e^2)^{1/2}}{e}(\dot{L}\mathbf{P} + \dot{M}\mathbf{Q}) + \dfrac{\cot i}{(1-e^2)^{1/2}}\dfrac{\partial \dot{\mathbf{x}}}{\partial i}\right]$	$\dfrac{\partial \omega}{\partial \dot{\mathbf{x}}} = \dfrac{-1}{na^2}\left[\dfrac{-(1-e^2)^{1/2}}{e}(L\mathbf{P} + M\mathbf{Q}) + \dfrac{\cot i}{(1-e^2)^{1/2}}\dfrac{\partial \mathbf{x}}{\partial i}\right]$
$\dfrac{\partial \Omega}{\partial \mathbf{x}} = \dfrac{-1}{na^2(1-e^2)^{1/2}\sin i}\dfrac{\partial \dot{\mathbf{x}}}{\partial i}$	$\dfrac{\partial \Omega}{\partial \dot{\mathbf{x}}} = \dfrac{+1}{na^2(1-e^2)^{1/2}\sin i}\dfrac{\partial \mathbf{x}}{\partial i}$

4. The Set-III Orbit Elements

We will now introduce a new set of orbit elements defined by a system of non-integrable differential relations. Two of the orbit elements will be left unchanged: a and e. The four other orbit elements (M_0, i, ω, Ω) are replaced by four new orbit elements (λ, r, p, q), according to the following definitions:

$$\left.\begin{aligned}\Delta\lambda &= \cos i\,\Delta\Omega + \Delta\omega + \Delta M_0 \\ \Delta r &= \cos i\,\Delta\Omega + \Delta\omega \\ \Delta p &= +\cos\omega\,\Delta i + \sin\omega\sin i\,\Delta\Omega \\ \Delta q &= -\sin\omega\,\Delta i + \cos\omega\sin i\,\Delta\Omega\end{aligned}\right\} \tag{15}$$

Solving these four relations for $\Delta i, \Delta\Omega, \Delta\omega, \Delta M_0$ gives

$$\left.\begin{aligned}\Delta i &= \cos\omega\,\Delta p - \sin\omega\,\Delta q \\ \Delta\Omega &= \frac{\sin\omega}{\sin i}\Delta p + \frac{\cos\omega}{\sin i}\Delta q \\ \Delta\omega &= -\sin\omega\cot i\,\Delta p - \cos\omega\cot i\,\Delta q + \Delta r \\ \Delta M_0 &= -\Delta r + \Delta\lambda\end{aligned}\right\} \tag{16}$$

We have thus defined the set-III orbit elements given by Brouwer and Clemence (Ref. 2, p. 241). We will now give an elementary derivation of the partial derivatives related to the set-III orbit elements. We derive these partial derivatives from the partial derivatives given in the preceding subsections for the classical orbit elements

$(a, e, M_0, i, \omega, \Omega)$. We will always use the set-III orbit elements in the order (a, e, λ, r, p, q). We will designate them symbolically by the letter p or p_α. According to the basic relations (Expression 16), the 36 partial derivatives of a_α with respect to p_β may be assembled in the following 6×6 matrix formed with four blocks of 3×3 matrices:

$$\left[\frac{\partial a_\alpha}{\partial p_\beta}\right] = \left[\begin{array}{c|c} I_3 & B \\ \hline 0 & D \end{array}\right] \tag{17}$$

where the matrices B and D are

$$B = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{vmatrix} \tag{18}$$

$$D = \begin{vmatrix} 0 & \cos\omega & -\sin\omega \\ +1 & \dfrac{-\sin\omega}{\operatorname{tg} i} & \dfrac{-\cos\omega}{\operatorname{tg} i} \\ 0 & \dfrac{\sin\omega}{\sin i} & \dfrac{\cos\omega}{\sin i} \end{vmatrix} \tag{19}$$

The matrix D is thus the matrix with partial derivatives of (i, ω, Ω) with respect to (r, p, q). The partial derivatives of $\mathbf{x}$ and $\dot{\mathbf{x}}$ with respect to the set-III orbit elements p_α are obtained by multiplying the Matrices (10) and (17):

$$\left[\frac{\partial \mathbf{x}}{\partial p_\beta}\right] = \left[\frac{\partial \mathbf{x}}{\partial a_\alpha}\right]\left[\frac{\partial a_\alpha}{\partial p_\beta}\right] \tag{20}$$

This matrix multiplication is straightforward because of the presence of many zeros and one's in the second matrix. Only the terms in the right half of the first matrix which combine with the D terms of the second matrix require some attention. These terms are in the following 3×3 matrix:

$$\left[\frac{\partial \mathbf{x}}{\partial r}\ \frac{\partial \mathbf{x}}{\partial p}\ \frac{\partial \mathbf{x}}{\partial q}\right] = \left\{X\left[\frac{\partial \mathbf{P}}{\partial i}\right] + Y\left[\frac{\partial \mathbf{Q}}{\partial i}\right]\right\} \cdot D \tag{21}$$

where the two 3×3 matrices $[\partial \mathbf{P}/\partial i]$ and $[\partial \mathbf{Q}/\partial i]$ at the right side have been defined in Expression (9). We have the two products:

$$\left.\begin{aligned} \left[\frac{\partial \mathbf{P}}{\partial i}\right] \cdot D &= [\mathbf{Q}, 0, -\mathbf{R}] \\ \left[\frac{\partial \mathbf{Q}}{\partial i}\right] \cdot D &= [-\mathbf{P}, \mathbf{R}, 0] \end{aligned}\right\} \tag{22}$$

Equation (20) may then be written in the form

$$\left[\frac{\partial \mathbf{x}}{\partial r}\ \frac{\partial \mathbf{x}}{\partial p}\ \frac{\partial \mathbf{x}}{\partial q}\right] = [\mathbf{Q}X - \mathbf{P}Y, +\mathbf{R}Y, -\mathbf{R}X] \tag{23}$$

The expression $\mathbf{Q}X - \mathbf{P}Y$ may be replaced by a cross product of vectors:

$$\mathbf{Q}X - \mathbf{P}Y = \mathbf{R} \times \mathbf{x} \tag{24}$$

The final Matrix (20) with the partial derivatives with respect to the set-III orbit elements may thus be written as follows:

$$\left[\frac{\partial \mathbf{x}}{\partial p_\alpha}\right] = \begin{array}{c|c|c|c|c|c} a & e & r & \lambda & p & q \\ \hline \dfrac{1}{a}\left(\mathbf{x} - \dfrac{3}{2}\dot{\mathbf{x}}t\right) & LP + MQ & -\dfrac{\dot{\mathbf{x}}}{n} + \mathbf{R} \times \mathbf{x} & \dfrac{\dot{\mathbf{x}}}{n} & \mathbf{R}Y & -\mathbf{R}X \\ \hline \dfrac{-1}{2a}\left(\dot{\mathbf{x}} - 3\mu\dfrac{\mathbf{x}t}{r^3}\right) & \dot{L}\mathbf{P} + \dot{M}\mathbf{Q} & n\left(\dfrac{a}{r}\right)^3 \mathbf{x} + \mathbf{R} \times \dot{\mathbf{x}} & -n\left(\dfrac{a}{r}\right)^3 \mathbf{x} & \mathbf{R}\dot{Y} & -\mathbf{R}\dot{X} \end{array} \tag{25}$$

5. Some Additional Formulas Related to the Set-III Orbit Elements

If one wants to compute planetary perturbations with some of the methods developed in SPS 37-49, Vol. III, pp. 31–40, with the set-III orbit elements, it is necessary to be in possession of the inverse partial derivatives $\partial p_\alpha/\partial \mathbf{x}$. These partial derivatives may be obtained from Matrix (25) by using the set-III formulas corresponding to Expression (13); this requires knowledge of the Poisson parentheses. Thus, let us develop the Lagrange brackets and Poisson parentheses for the set-III orbit elements p_α. These quantities may be considered as two skew-symmetric tensors of rank 2 in a six-dimensional space. The transformation from the six elements a_λ to the six new elements p_α may thus be expressed by the classical formulas of transformation of rank-2 tensors.

$$\left.\begin{aligned} [p_\alpha, p_\beta] &= \sum_{\lambda,\mu} [a_\lambda, a_\mu] \frac{\partial a_\lambda}{\partial p_\alpha}\frac{\partial a_\mu}{\partial p_\beta} \\ (p_\alpha, p_\beta) &= \sum_{\lambda,\mu} (a_\lambda, a_\mu) \frac{\partial p_\alpha}{\partial a_\lambda}\frac{\partial p_\beta}{\partial a_\mu} \end{aligned}\right\} \tag{26}$$

Expression (26) would be easier to apply in a matrix form:

$$\left.\begin{aligned}[[p_\alpha, p_\beta]] &= \left[\frac{\partial a_\lambda}{\partial a_\alpha}\right]^T \left[[a_\lambda, a_\mu]\right]\left[\frac{\partial a_\mu}{\partial p_\beta}\right] \\ [(p_\alpha, p_\beta)] &= \left[\frac{\partial p_\alpha}{\partial a_\lambda}\right][(a_\lambda, a_\mu)]\left[\frac{\partial p_\beta}{\partial a_\mu}\right]^T\end{aligned}\right\} \tag{27}$$

We have used Expression (27) to obtain the set-III Poisson parentheses and Lagrange brackets. We find that for each case there are only four pairs of non-zero elements, all the other terms of the 6×6 matrices being zero. The four basic Lagrange brackets are

$$\left.\begin{aligned}[a, r] &= \frac{na}{2}\left[1 - (1 - e^2)^{1/2}\right] \\ [a, \lambda] &= \frac{-na}{2} \\ [e, r] &= \frac{na^2 e}{(1 - e^2)^{1/2}} \\ [p, q] &= na^2 (1 - e^2)^{1/2}\end{aligned}\right\} \tag{28}$$

The four Poisson parentheses are

$$\left.\begin{aligned}(a, \lambda) &= \frac{-2}{na} \\ (e, r) &= \frac{(1 - e^2)^{1/2}}{na^2 e} \\ (e, \lambda) &= \frac{(1 - e^2)^{1/2}}{na^2 e}\left[1 - (1 - e^2)^{1/2}\right] \\ &= \frac{e(1 - e^2)^{1/2}}{na^2\left[1 + (1 - e^2)^{1/2}\right]} \\ (p, q) &= \frac{1}{na^2 (1 - e^2)^{1/2}}\end{aligned}\right\} \tag{29}$$

The partial derivatives of the set-III orbit elements with respect to the position and velocity components may now be obtained by using Expression (13). With this formula no numerical difficulties with zero inclination will occur; but on the other hand, zero eccentricities may give rise to some difficulties, since the factor e appears in some denominators.

6. Some Elements Related to the Set-III Elements

It is of interest to introduce a set of elements composed of the three classical elements (a, e, M_0) and the three set-III elements (r, p, q). In this set of elements the mean anomaly M_0 is used instead of λ. This will be useful in the introduction of the set-I elements in *Subsection 7*.

The partial derivatives of the classical elements $a_\alpha = (a, e, M_0, i, \omega, \Omega)$ with respect to the elements $p_\beta = (a, e, M_0, r, p, q)$ are given in the following 6×6 matrix:

$$\left[\frac{\partial a_\alpha}{\partial p_\beta}\right] = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & D \end{array}\right] \tag{30}$$

where the 3×3 matrix D has been defined in Matrix (19). Matrix (30) is the analogue of Matrix (17), but we see that Matrix (30) is slightly more symmetric because B is replaced by zero. The consequence of this additional symmetry will be that the final partial derivatives will have a more simple form. The matrix products given in Expression (22) are still to be used here to form the partial derivatives of the coordinates $\mathbf{x}$. Matrix (23) will be the right part of the new matrix with partial derivatives. The left part, corresponding to the classical orbit elements (a, e, M_0) will be taken from Matrix (10). The complete matrix with the 36 partial derivatives of $\mathbf{x}$ and $\dot{\mathbf{x}}$ will then take the form:

$$\left[\frac{\partial \mathbf{x}}{\partial p_\alpha}\right] = \begin{array}{c|cccccc} & a & e & M_0 & r & p & q \\ \hline & \dfrac{1}{a}\left(\mathbf{x} - \dfrac{3}{2}\dot{\mathbf{x}}t\right) & LP + MQ & \dfrac{\dot{\mathbf{x}}}{n} & QX - PY & RY & -RX \\ & \dfrac{-1}{2a}\left(\dot{\mathbf{x}} - 3\mu\dfrac{\mathbf{x}t}{r^3}\right) & \dot{L}P + \dot{M}Q & -n\left(\dfrac{a}{r}\right)^3\mathbf{x} & Q\dot{X} - P\dot{Y} & R\dot{Y} & -R\dot{X} \end{array} \tag{31}$$

We can see here that the only column of this matrix which is different than the columns in Matrix (25) is the column corresponding to the variable r.

The fundamental Lagrange brackets and Poisson parentheses of these orbit elements are

$$\left.\begin{aligned} [a, M_0] &= -\frac{na}{2}, & (a, M_0) &= \frac{-2}{na} \\ [a, r] &= -\frac{na}{2}(1-e^2)^{1/2}, & (e, M_0) &= -\frac{(1-e^2)}{na^2e} \\ [e, r] &= \frac{na^2e}{(1-e^2)^{1/2}}, & (e, r) &= \frac{(1-e^2)^{1/2}}{na^2e} \\ [p, q] &= na^2(1-e^2)^{1/2}, & (p, q) &= \frac{1}{na^2(1-e^2)^{1/2}} \end{aligned}\right\} \tag{32}$$

7. The Set-I Orbit Elements

We will derive from the preceding results a set of orbit elements which has been called set-I by Brouwer and Clemence (Ref. 2, p. 238). They are related to the set-III elements in the following way: the first three elements (a, e, M_0) which are used are the classical elements. The three other elements (ψ_1, ψ_2, ψ_3) will replace the set-III elements p, q, and r, according to the following differential relations:

$$\left.\begin{aligned} \Delta\psi_1 &= P_x\,\Delta p + Q_x\,\Delta q + R_x\,\Delta r \\ \Delta\psi_2 &= P_y\,\Delta p + Q_y\,\Delta q + R_y\,\Delta r \\ \Delta\psi_3 &= P_z\,\Delta p + Q_z\,\Delta q + R_z\,\Delta r \end{aligned}\right\} \tag{33}$$

The nine partial derivatives of (ψ_1, ψ_2, ψ_3) with respect to the three elements (p, q, r) are

$$\left[\frac{\partial\psi}{\partial p}\right] = \begin{bmatrix} P_x & Q_x & R_x \\ P_y & Q_y & R_y \\ P_z & Q_z & R_z \end{bmatrix} \tag{34}$$

and the nine partial derivatives of (p, q, r) with respect to the three elements (ψ_1, ψ_2, ψ_3) are obtained by taking the inverse (= transpose) of Matrix (34):

$$\left[\frac{\partial p}{\partial\psi}\right] = \begin{bmatrix} P_x & P_y & P_z \\ Q_x & Q_y & Q_z \\ R_x & R_y & R_z \end{bmatrix} \tag{35}$$

We will now be able to obtain the partial derivatives of $(x, y, z, \dot{x}, \dot{y}, \dot{z})$ with respect to $(a, e, M_0, \psi_1, \psi_2, \psi_3)$ by simple matrix multiplications. The derivatives with respect to a, e, M_0 have been given previously. The derivatives with respect to (ψ_1, ψ_2, ψ_3) are obtained by multiplying Matrix (23) by Matrix (34); we find:

$$\left[\frac{\partial\mathbf{x}}{\partial\psi}\right] = \left[\frac{\partial\mathbf{x}}{\partial p}\right]\cdot\left[\frac{\partial p}{\partial\psi}\right] = \begin{bmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{bmatrix} \tag{36}$$

The full 6×6 matrix with the partials with respect to $(a, e, M_0, \psi_1, \psi_2, \psi_3)$ is

$$\left[\frac{\partial\mathbf{x}}{\partial\psi}\right] = \begin{array}{c|ccc|ccc} & a & e & n_0 & \psi_1 & \psi_2 & \psi_3 \\ \hline & \multirow{3}{*}{$\frac{1}{a}\left(\mathbf{x} - \frac{3}{2}\dot{\mathbf{x}}t\right)$} & \multirow{3}{*}{$L\mathbf{P} + M\mathbf{Q}$} & \multirow{3}{*}{$\frac{\dot{\mathbf{x}}}{n}$} & 0 & z & $-y$ \\ & & & & $-z$ & 0 & x \\ & & & & y & $-x$ & 0 \\ \hline & \multirow{3}{*}{$\frac{-1}{2a}\left(\dot{\mathbf{x}} - 3\mu\frac{\mathbf{x}}{r^3}t\right)$} & \multirow{3}{*}{$\dot{L}\mathbf{P} + \dot{M}\mathbf{Q}$} & \multirow{3}{*}{$-n\left(\frac{a}{r}\right)^3\mathbf{x}$} & 0 & $\dot{z}$ & $-\dot{y}$ \\ & & & & $-\dot{z}$ & 0 & $\dot{x}$ \\ & & & & $\dot{y}$ & $-\dot{x}$ & 0 \end{array} \tag{37}$$

Thus, we find here the results obtained by Brouwer and Clemence (Ref. 2, p. 238). These are also essentially the results which have recently been obtained by D. C. Lewis (Ref. 3, p. 107). These authors obtained the same results

in a different way, essentially because they used the fact that (ψ_1, ψ_2, ψ_3) correspond to small rotations about the x, y, and z axes. We have not used this property here.

Matrix (37) gives a simple set of partial derivatives for the two-body problem. However, the inverse of Matrix (37) is somewhat more difficult to obtain than in the preceding cases, mainly because the Lagrange brackets and Poisson parentheses are more complicated here. We found, for instance, that some of the Lagrange brackets involving (ψ_1, ψ_2, ψ_3) are proportional to the components of the angular momentum vector. More precisely, we find that 20 out of the 36 Lagrange brackets are non-zero. The 10 fundamental brackets are:

$$\left.\begin{aligned}
[a, M_0] &= -\frac{na}{2} \\
[\psi_1, \psi_2] &= na^2(1-e^2)^{1/2}\cos i \\
[\psi_1, \psi_3] &= na^2(1-e^2)^{1/2}\cos\Omega\sin i \\
[\psi_2, \psi_3] &= na^2(1-e^2)^{1/2}\sin\Omega\sin i \\
[a, \psi_1] &= -\frac{na}{2}(1-e^2)^{1/2}\sin\Omega\sin i \\
[a, \psi_2] &= +\frac{na}{2}(1-e^2)^{1/2}\cos\Omega\sin i \\
[a, \psi_3] &= -\frac{na}{2}(1-e^2)^{1/2}\cos i \\
[e, \psi_1] &= +na^2e\sin\Omega\frac{\sin i}{(1-e^2)^{1/2}} \\
[e, \psi_2] &= -na^2e\cos\Omega\frac{\sin i}{(1-e^2)^{1/2}} \\
[e, \psi_3] &= +na^2e\frac{\cos i}{(1-e^2)^{1/2}}
\end{aligned}\right\} \quad (38)$$

The Poisson parentheses can be found by inversion of the matrix with Lagrange brackets. We find that there are eight basic non-zero Poisson parentheses:

$$\left.\begin{aligned}
(a, M_0) &= -\frac{2}{na} \\
(e, M_0) &= \frac{-(1-e^2)}{na^2e} \\
(e, \psi_1) &= \frac{+(1-e^2)^{1/2}}{na^2e}\sin i\sin\Omega \\
(e, \psi_2) &= \frac{-(1-e^2)^{1/2}}{na^2e}\sin i\cos\Omega \\
(e, \psi_3) &= \frac{+(1-e^2)^{1/2}}{na^2e}\cos i \\
(\psi_1, \psi_2) &= \frac{\cos i}{na^2(1-e^2)^{1/2}} \\
(\psi_1, \psi_3) &= \frac{\sin i\cos\Omega}{na^2(1-e^2)^{1/2}} \\
(\psi_2, \psi_3) &= \frac{\sin i\sin\Omega}{na^2(1-e^2)^{1/2}}
\end{aligned}\right\} \quad (39)$$

Having all the Poisson parentheses, we can now easily construct the matrix of inverse partial derivatives of $(a, e, M_0, \psi_1, \psi_2, \psi_3)$ with respect to $(\mathbf{x}, \dot{\mathbf{x}})$ by simple matrix multiplications.

References

1. Danby, J. M. A., "Integration of the Equation of Planetary Motion in Rectangular Coordinates," *Astron. J.*, Vol. 67, No. 5, pp. 287–299, June 1962.

2. Brouwer, D., and Clemence, G., *Methods of Celestial Mechanics*, Academic Press, New York, 1961.

3. Lewis, D. C., "Group Theoretical Aspects of the Perturbation of Keplerian Motion," NASA PM-67-21. NASA Electronics Research Center, Cambridge, Massachusetts, January 1968.

II. Computation and Analysis

SYSTEMS DIVISION

A. LEASTQ: A Program for Least-Squares Fitting Segmented Cubic Polynomials Having a Continuous First Derivative, *T. M. Lang*

1. Introduction

Although there are several existing subroutines for linear least-squares curve fitting, the Fortran IV program LEASTQ offers a number of important advantages. At the heart of the procedure is a subroutine SPLHFT[1] which applies the highly successful Householder reflection technique[2] to least-square solve the resulting system of parameterized equations. In particular, the block structure of this least-squares problem allows significant savings of computer storage and execution time that has not been previously realized.

The complete LEASTQ package includes other optional features. If desired, a computer-generated Fortran IV subprogram can be requested to evaluate the fitted function and its derivative. (The break points and polynomial coefficients are coded in DATA statements.) Another set of subroutines plots the data and fitted function on convenient and readable graphs.

2. Mathematical Method

Let (x_i,y_i), $i = 1, \cdots, n$, denote the coordinates of the observations, reordered such that $x_1 \leq x_2 \leq \cdots \leq x_n$. Consider the $k + 1$ points $a = b_1 < b_2 < \cdots < b_{k+1} = c$ with $a \leq x_1 < x_n \leq c$. Although it is possible to solve directly for the coefficients of k cubic polynomials where the first derivative is continuous on $[a, c]$, there exists a parameterization to the problem

$$\text{minimize} \left| \left| Ax - b \right| \right| \text{ subject to } Cx = d$$

in which the constraints $Cx = d$ are implicit, and the banded nature of the coefficient matrix becomes apparent.[3]

Suppose an $x_i \in [b_j, b_{j+1}]$ for some integer j, $1 \leq j \leq k$. Compute $r = b_{j+1} - b_j$ and $s = (x_i - b_j)/r$. The standard

[1]Lang, T. M., and Hanson, R. J., *LEASTQ—A Least Squares Package Using C^1 Segmented Cubic Polynomials*, JPL internal document (in process).

[2]Hanson, R. J., and Lawson, C. L., *Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems, Part I: Extensions*, JPL internal document, July 12, 1968, pp. 28–32.

[3]Lawson, C. L., *Least Squares Curve Fitting Using Segmented Quadratic or Cubic Polynomials Having C^1 Continuity*, Jan. 24, 1968 (JPL internal document).

Hermite cubic interpolation formula

$$f(x_i) = f(b_j)p_1(x_i) + f'(b_j)p_2(x_i) + f(b_{j+1})p_3(x_i) + f'(b_{j+1})p_4(x_i)$$

where

$$p_1(x_i) = (1 + 2s)(s - 1)^2, \qquad p_2(x_i) = rs(s - 1)^2$$
$$p_3(x_i) = s^2(3 - 2s), \qquad p_4(x_i) = rs^2(s - 1)$$

constitutes a one-to-one linear mapping with the original least-squares problem when we solve for the $2k + 2$ parameters $(f(b_j), f'(b_j))$, $j = 1, \cdots, k+1$, to minimize

$$\left[\sum_{i=1}^{n} (y_i - f(x_i))^2\right]^{1/2}$$

Each subinterval $[b_j, b_{j+1}]$, $j = 1, \cdots, k$, yields a block of equations, one equation for each $x_i \in [b_j, b_{j+1}]$. Because of the constraints on $f(b_j)$ $f'(b_j)$, they appear in the equations for $[b_{j-1}, b_j]$ and $[b_j, b_{j+1}]$, but no other. Hence, each block consists of exactly 4 nonzero column entries and is offset two columns to the right with respect to the previous block (Fig. 1).

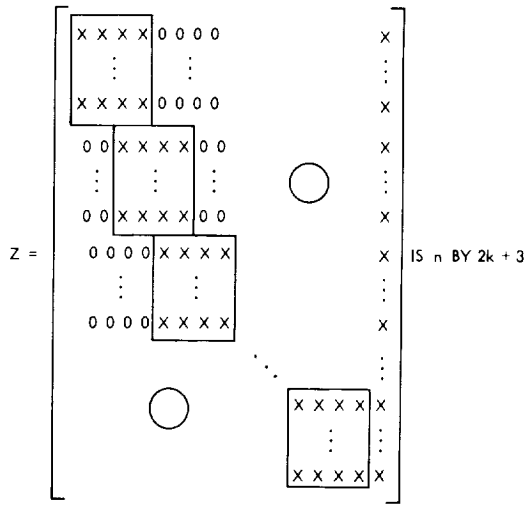


Fig. 1. Block matrix of parameterized coefficients

3. Numerical Procedure

The system of parameterized equations, if solved in a single step normally would require a matrix Z with n rows and $2k + 3$ columns, although many entries are zeroes. In fact, for each row indexed i, $i = 1, \cdots, n$, the only nonzero entries in Z are

$$z_{i,2j-1} = p_1(x_i), \qquad z_{i,2j+2} = p_4(x_i)$$
$$z_{i,2j+1} = p_3(x_i), \qquad z_{i,2k+3} = y_i$$
$$z_{i,2j} = p_2(x_i)$$

for $x \in [b_j, b_{j+1}]$. An existing program SEQLSQ[4] can process the blocks sequentially or accept the entire matrix, though the number of columns is fixed by the number of subintervals $[b_j, b_{j+1}]$, $j = 1, \cdots, k$. A problem with 500 data points and 32 subintervals ultimately would need $500 \times (2 \times 32 + 3) = 500 \times 67 = 33{,}500$ core locations, which exceeds the present IBM 7094 capacity.

LEASTQ and its subroutine SPLHFT are designed to take advantage of this block structure. The blocks can be processed sequentially in such a way that only 5 columns are needed in the working array.

The parameter equations for the first subinterval are written into a matrix, e.g., W, in columns 1 through 5. Householder orthogonal transformations are applied to reduce the entries to upper-triangular form, and then the third and fourth rows are rearranged to prepare for the next block (Fig. 2). The single entry in row 5 is the block residual term.

The next block of equations begins in row 5 and the transformations are applied to rows 3, 4 and below to again yield an upper-triangular array (Fig. 3). This procedure is continued until all k blocks have been processed.

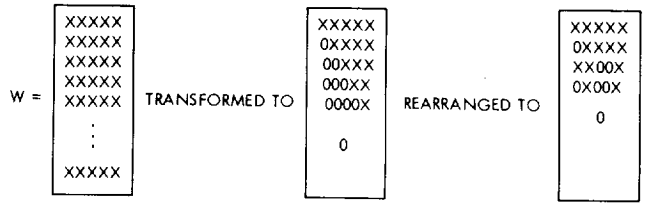


Fig. 2. First iteration

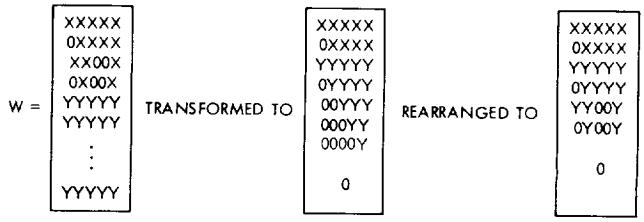


Fig. 3. Second iteration

[4]Hanson, R. J., *Write-ups for LSQSOL, COVLSQ, LSQLS2, and SEQSLQ/SEQLQ2*, April 24, 1968, pp. 21–29 (JPL internal document).

Then a back-substitution step yields the $2k + 2$ parameters $(f(b_j), f'(b_j)), j = 1, \cdots, k+1$.

Very few of the zeroes in the Z matrix enter calculations in this arrangement of the array W. Column requirements are reduced from $2k + 3$ to 5, and the worst-case row reduction is from n to ($2k$ + maximum points in any subinterval). This procedure, however, requires that at least 4 data points must fall in the first subinterval and at least 2 points in each of the other subintervals to ensure the existence of a unique solution. If our example of 500 points and 32 subintervals had no more than 20 points in any subinterval, $5 \times (2 \times 32 + 20) = 5 \times 84 = 420$ core locations would suffice, and execution time would be significantly decreased.

4. Remarks

The computations in LEASTQ are performed in double precision, and solutions to test problems agree precisely with those from SEQLSQ. A test case with 3 break points and 20 data points was executed in 0.20 sec using SEQLSQ, but only 0.10 sec was needed by LEASTQ. With LEASTQ, execution time is a linear function of both the number of data points and the number of break points.

The plotting section is independent of solution to the least-squares problem, thereby eliminating unneeded coding when this option is not used. All data and the fitted function are scaled together, then plotted using a modification of EZPLOT.[5] Arbitrarily, 300 or fewer data points appear per frame until all points are plotted. (The number 300 can be reset by the user.) Scaling is identical on consecutive frames. For each frame of data and fitted function, a second frame displays the continuous first derivative, with the same abscissa. Break points are indicated at the bottom of each graph.

Another optional section will generate a Fortran IV subprogram deck for the user. With two entry points, the punched deck will evaluate $f(x)$ or $f'(x)$ in double precision, when the user supplies the argument x. Break points and parameter coefficients are coded into DATA statements.

5. Utility

LEASTQ has already been used successfully to smooth data. Constraining the first derivative ensures a "rather nice" fit of the points. Execution is very rapid. One large test case of 1000 data points and 25 break points required only 2.8 sec.

There is no present provision to alter or add break points to reduce the residual. However, the least-squares fit is sensitive to the choice of break points, and modifications of this type can have unsuspected results on the residual and the continuous derivative.

Another possible extension is the use of fifth-degree segmented polynomials, instead of cubics, with both the first and second derivatives being continuous. Such a parameterization might make more sense if break-point juggling is contemplated. Some users are more interested in the first derivative than in the polynomial itself, and the cubic approach makes such data smoothing more difficult.

6. Displays

Figure 4 is a sample of the Fortran IV subprogram which is punched out for the user to evaluate the fitted function and its derivative. The constants in the DATA statements are the break points and parameterized polynomial coefficients for each subinterval. An entry to FXDP with the argument X will initiate a search for the proper subinterval containing X, then evaluate the polynomial X. An entry to DYDXDP performs the same search and computes the derivative at X. Extrapolation occurs if the desired point is not within one of the subintervals.

Segmented cubic polynomials were applied successfully to smooth discrete doppler radar data from *Lunar Orbiter* V and the derivative was used to obtain the first gravimetric map of the front side of the moon. Figure 5 is a frame showing the data from one orbit in its least-squares polynomial fit, and Fig. 6 shows the continuous first derivative (Ref. 1).

Data points are denoted with an asterisk, and break points are shown as an "X" along the lower edge of the graph. The user supplies the title line for each frame. In terms of execution time, plotting is the most expensive option: about 5 sec were needed to produce these two frames.

[5]A JPL Sect. 314 program.

Reference

1. Muller, T., and Sjogren, W. R., *Consistency of Lunar Orbiter Residuals With Trajectory and Local Gravity Effects*, Technical Report 32-1307. Jet Propulsion Laboratory, Pasadena, Calif., September 1, 1968.

```
$IBFTC .PARA.  LIST
C
C       SAMPLE COMMENT.
C
C       Y,M,D,H,M=68,07,24,16,31 7094A
C
        DOUBLE PRECISION FUNCTION FXDP  (X)
        REAL B(10)
        DOUBLE PRECISION X,Z( 20),R,Y,Q1,Q2,Q3,Q4,Q5
        DATA NB/10/,LOOK/1/
        DATA B( 1),B( 2) / 0.0000000E-39, 1.6000000E 01/
        DATA B( 3),B( 4) / 2.8000000E 01, 3.7000000E 01/
        DATA B( 5),B( 6) / 4.2000000E 01, 4.9000000E 01/
        DATA B( 7),B( 8) / 5.4000000E 01, 6.2000000E 01/
        DATA B( 9),B(10) / 7.4000000E 01, 9.0000000E 01/
        DATA Z(  1),Z(  2) / 9.561539580380041D-02, 1.128145129085036D-01/
        DATA Z(  3),Z(  4) / 9.501712619355640D-02,-8.160088780949285D-02/
        DATA Z(  5),Z(  6) /-1.708740185881998D-01,-2.373248249413888D-02/
        DATA Z(  7),Z(  8) / 2.407657424848087D-01, 1.158097622495027D-02/
        DATA Z(  9),Z( 10) / 1.769927521742830D 00, 1.140639342263009D-01/
        DATA Z( 11),Z( 12) / 5.049514670611926D-01,-5.912552096700325D-02/
        DATA Z( 13),Z( 14) / 3.574101493610389D-01, 2.968269119769930D-01/
        DATA Z( 15),Z( 16) / 1.144302401891606D 00,-4.482827378377042D-03/
        DATA Z( 17),Z( 18) / 2.165028826547599D-01,-7.650635429082006D-02/
        DATA Z( 19),Z( 20) / 4.395104844602346D-01, 1.882921971235911D-01/
        INTJ=1
        GO TO 5
        ENTRY DYDXDP(X)
        INTJ=2
   5    IF(X.LT.DBLE(B( 9))) GO TO 10
        LOOK= 9
        GO TO 20
  10    IF(X.LT.DBLE(B(LOOK))) LOOK=1
  15    IF(X.LE.DBLE(B(LOOK+1))) GO TO 20
        LOOK=LOOK+1
        GO TO 15
  20    R=DBLE(B(LOOK+1)-B(LOOK))
        Y=(X-DBLE(B(LOOK)))/R
        Q1=Y-1.D0
        J=2*LOOK-1
        GO TO (30,40),INTJ
  30    Q4=2.D0*Y
        Q2=1.D0+Q4
        Q3=3.D0-Q4
        Q4=Q1**2
        Q5=Y**2
        FXDP   =Q4*(Q2*Z(J)+R*Y*Z(J+1))+Q5*(Q3*Z(J+2)+R*Q1*Z(J+3))
        RETURN
  40    Q2=3.D0*Y
        Q3=6.D0*Y/R
        FXDP   =Q1*(Q3*(Z(J)-Z(J+2))+(Q2-1.D0)*Z(J+1))+Y*(Q2-2.D0)*Z(J+3)
        RETURN
        END
```

Fig. 4. Sample evaluation subprogram

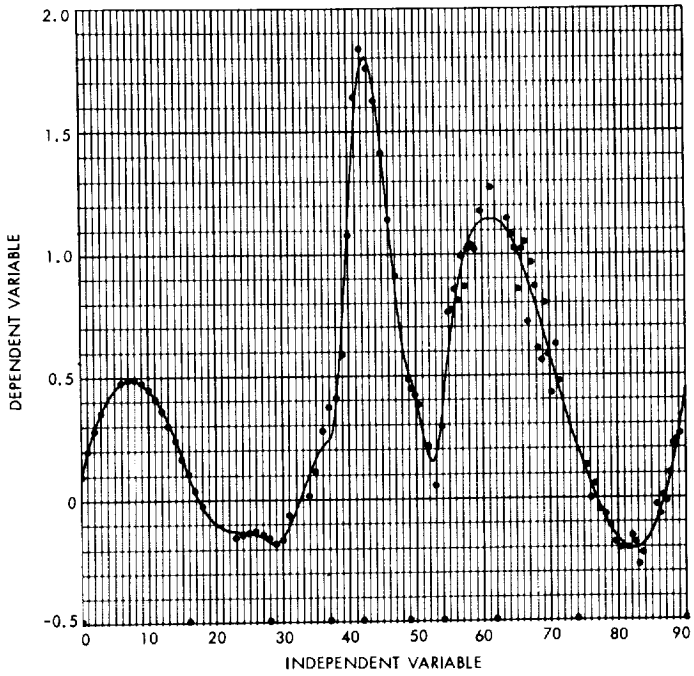


Fig. 5. *Lunar Orbiter V* velocity data and fit

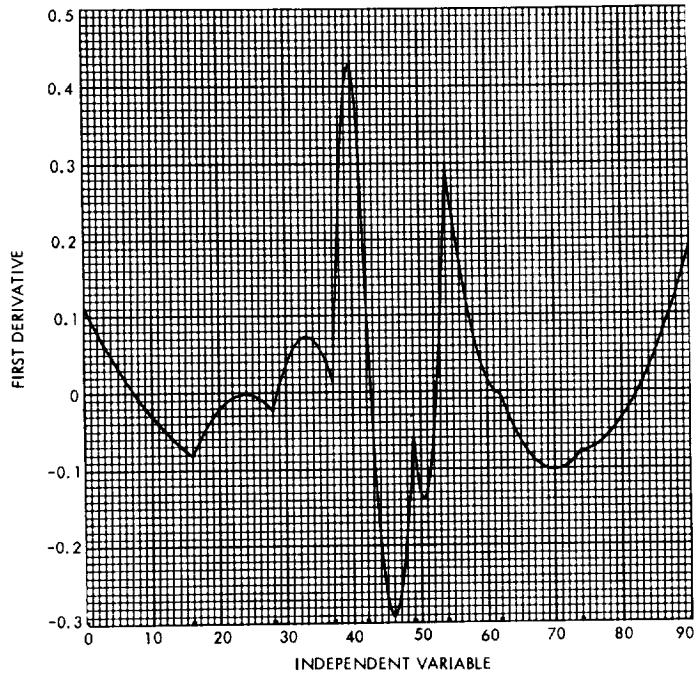


Fig. 6. Continuous first derivative

B. Abstracts of Certain Mathematical Subroutines, III: Minimization Subject to Linear Inequality and Equality Constraints With an Application to Curve Fitting, *R. J. Hanson and A. J. Semtner*

Abstracts of several general-purpose mathematical subroutines which have been developed or extended by JPL since September 1966 were reported in SPS 37-48, Vol. III, p. 26, and in SPS 37-50, Vol. III, p. 66. This article presents the following three abstracts:

(1) A general-purpose highly reliable subroutine is formulated which will minimize certain classes of real valued functions of several variables with linear inequality and equality constraints on these variables. The algorithm used is the gradient projection method in Ref. 1.

(2) The linear least-squares problem with these linear constraints is reformulated in such a way that the solution can be easily obtained with the gradient projection algorithm mentioned above.

(3) The subroutine in abstract 2 is applied to a problem of curve-fitting experimental data by so-called "spline functions" (continuous, twice differentiable jointed cubic polynomials in one independent variable) with the constraints that the first derivative is to be decreasing on the entire interval, while the function is to be convex downward on part of the interval and convex upward on the remaining part of the interval.

There are many other possible applications which one could make with these subroutines, including corrections to certain planetary constants past only a fixed number of decimal places, and the optimal allocation of resources from mathematical economic theory.

Abstract 1[a]

a. Purpose. Let $f(\mathbf{x})$ be a real-valued function of n variables $\mathbf{x} = (x_1, \cdots, x_n)$, which has continuous partial derivatives $(D_j f)(\mathbf{x})$ on $C = \{\mathbf{x} \mid C\mathbf{x} \geq \mathbf{c}\}$. Here C is a given $m_1 \times n$ real matrix and $\mathbf{c}$ is a real m_1 vector.

Then if f is convex, the gradient projection algorithm will find $\bar{\mathbf{x}} \in C$ (if it exists) such that $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$, $(\mathbf{x} \in C)$. Recall that if f has second partial derivatives, then f is convex in C if, and only if, its Hessian matrix $H = \{(D_i(D_j f)(\mathbf{x})\}$ is positive semidefinite in C (Ref. 2).

[a]Identification: subroutine name—GPMTHD/GPMTH2; source language—FORTRAN IV; machine—IBM 7094.

The subroutine may be used formally even if f is not convex on C, and will only allow the value of f to decrease in any case.

The user is required to supply code which will evaluate $f(\mathbf{x})$ and $\{(D_j f)(\mathbf{x})\}$ for any $\mathbf{x} \in C$, but the program itself "reaches" C by an internal algorithm which we will not elaborate upon here.

b. Mathematical method. J. B. Rosen's gradient projection algorithm is used to minimize f on C, provided this minimum exists (Ref. 1).

The main computational problem in this algorithm is the solution of a sequence of least-squares problems

$$A_i y_i \cong \{(D_j f)(x_i)\}, \qquad (i = 1, 2, \cdots)$$

The matrices A_i are of dimension $m_i \times n_i$, where m_i and n_i are positive integers. Thus it is necessary to have a reliable subroutine for solving such least-squares problems. Such a subroutine is provided by a slightly modified version of LSQSOL, previously described in SPS 37-48, Vol. III.

This subroutine also permits one to minimize f subject to equality constraints

$$E\mathbf{x} = \mathbf{e} \tag{1}$$

For in one case $r = \text{rank}(E) < n$, one can find the most general solution of Eq. (1) in the form

$$\mathbf{x} = \mathbf{x}_0 + H\mathbf{y} \tag{2}$$

where $\mathbf{x}_0 = E^{+}\mathbf{e}$, $EH = 0$, and $H^T H = I_{n-r}$.

If $E\mathbf{x}_0 = \mathbf{e}$, the problem then reduces to minimizing $g(\mathbf{y}) = f(\mathbf{x}_0 + H\mathbf{y})$ with gradient vector

$$\{(D_j g)(\mathbf{y})\} = \{(D_j f)(\mathbf{x}_0 + H\mathbf{y})\}H \tag{3}$$

in the region

$$C_H = \{\mathbf{y} \mid CH\mathbf{y} \geq \mathbf{c} - C\mathbf{x}_0\} \tag{4}$$

The program will also handle special types of bounds on the variables (such as nonnegativity of the variables) with a great savings in storage.

c. Experience. This subroutine has been tested and performed satisfactorily on several problems, and seems to be free of coding errors. Its use is straightforward.

The authors recommend that minimization problems with linear constraints be attempted with this subroutine.

d. Special class of minimization problem. For an arbitrary $m_A \times n_A$ matrix $A = \{a_{ij}\}$, and an m_A-vector $\mathbf{b}$, find $\tilde{\mathbf{x}}$ which minimizes the euclidean norm of

$$A\mathbf{x} - \mathbf{b} \tag{5}$$

subject to $m_E \leq n_A$ equality constraints

$$E\mathbf{x} = \mathbf{e} \tag{6}$$

and m_C inequality constraints

$$C\mathbf{x} \geq \mathbf{c} \tag{7}$$

Here E and C are, respectively, $m_E \times n_A$ and $m_C \times n_A$ matrices, while $\mathbf{e}$ and $\mathbf{c}$ are m_E and m_C vectors.

Abstract 2[7]

a. Purpose. Provide a general-purpose subroutine which solves the problem of Expression (5) subject to Eq. (6) and Expression (7).

b. Method. Equation (6) is first solved to yield an n_A vector $\mathbf{x}_0$ and an $n_A \times (n_A - m_E)$orthonormal matrix H such that

$$\mathbf{x} = \mathbf{x}_0 + H\mathbf{y} \tag{8}$$

is the most general solution of Eq. (6). The vector $\mathbf{y}$ is still arbitrary. In case rank $(E) = r_E < m_E$, or $r_E = 0$, program execution stops.

With

$$D = AH \tag{9}$$

$$\mathbf{d} = \mathbf{b} - A\mathbf{x}_0 \tag{10}$$

$$F = CH \tag{11}$$

$$\mathbf{f} = \mathbf{c} - C\mathbf{x}_0 \tag{12}$$

and

$$n_D = n_A - m_E, \qquad (n_D \leq m_A) \tag{13}$$

the original problem expressed in Eqs. (5), (6), and (7) is reduced to minimizing the euclidean norm of

$$D\mathbf{y} - \mathbf{d} \tag{14}$$

subject to the m_C inequality constraints

$$F\mathbf{y} \geq \mathbf{f} \tag{15}$$

The matrices D and F are, respectively, $m_A \times n_D$ and $m_C \times n_D$, while $\mathbf{d}$ and $\mathbf{f}$ are m_A- and m_C-vectors. The problem of Eqs. (14) and (15) is unchanged if we choose an $m_A \times m_A$ orthonormal matrix Q (Ref. 3) such that

$$Q(D\mathbf{y} - \mathbf{d}) = \begin{bmatrix} R \\ 0 \end{bmatrix} \mathbf{y} - \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} \begin{matrix} \} n_D \\ \} m_A - n_D \end{matrix} \tag{16}$$

where R is $n_D \times n_D$ and upper triangular.

The euclidean norm of the right member of Eq. (16) is minimized precisely when we minimize the euclidean norm of

$$R\mathbf{y} - \mathbf{g}_1 \tag{17}$$

subject to the inequality constraints

$$F\mathbf{y} \geq \mathbf{f} \tag{18}$$

If R is singular, the execution stops. If R is nonsingular, we let $R\mathbf{y} = \mathbf{z}$ and $\mathbf{z} - \mathbf{g}_1 = \mathbf{w}$ to obtain the equivalent problem of minimizing the function

$$||\mathbf{w}||^2/2 \tag{19}$$

subject to the inequality constraints

$$FR^{-1}\mathbf{w} \geq \mathbf{f} - FR^{-1}\mathbf{g}_1 \tag{20}$$

Once $\tilde{\mathbf{w}}$ is found which minimizes Ex. (19), subject to the inequalities of (20), we compute

$$\tilde{\mathbf{z}} = \tilde{\mathbf{w}} + \mathbf{g}_1 \tag{21}$$

$$\tilde{\mathbf{y}} = R^{-1}\tilde{\mathbf{z}} \tag{22}$$

and

$$\tilde{\mathbf{x}} = \mathbf{x}_0 + H\tilde{\mathbf{y}} \tag{23}$$

[7]Identification: subroutine name—LWLCV1; source language—FORTRAN IV; machine—IBM 7094.

The vector $\tilde{\mathbf{w}}$ is obtained with the gradient projection subroutine GPMTHD/GPMTH2.

c. Experience. Computational experience with LWLCV1 shows that it is a working subroutine which appears to be free of coding errors. It is, therefore, highly recommended.

Abstract 3, an Application of the Subroutine LWLCV1 to Curve-Fitting Experimental Data

Let k real pairs of points (x_i,y_i), $(i = 1, \cdots, k)$, be obtained as the result of some experiment taken on a bounded interval $[a,b]$, $(a < b)$. (Without loss of generality we may take $a = -1$, and $b = +1$).

Let $n_b \geq 0$ points $\hat{x}_i$, $(i = 1, \cdots, n_b)$, be chosen in the interior of $[-1,1]$. We then wish to construct a function $y = f(x)$, where f is defined piece-wise as a cubic polynomial on each of $n_b + 1$ subintervals. Thus if

$$I_i = \begin{cases} [-1,\hat{x}_1], & i = 1 \\ [\hat{x}_{i-1},\hat{x}_i], & 1 < i \leq n_b \\ [\hat{x}_{n_b},1], & i = n_b + 1 \end{cases} \tag{24}$$

then for $x \in I_i$

$$f(x) = a_{1i}x^3 + a_{2i}x^2 + a_{3i}x + a_{4i}, \qquad (i = 1, \cdots, n_b + 1) \tag{25}$$

We further require that $f, f' = df/dx$, and $f'' = d^2f/dx^2$ be continuous on $[-1,1]$. Let

$$\hat{\mathbf{a}} = [a_{11}, a_{21}, a_{31}, a_{41}, \cdots, a_{1,n_b+1}, \cdots, a_{4,n_b+1}]^T \tag{26}$$

be a $4(n_b + 1)$-vector to be determined.

We require that f be monotone decreasing on $[-1,1]$, and that f be convex downward on $[-1,\hat{x}_m]$ and convex upward on $[\hat{x}_m,1]$. Here $\hat{x}_m$ is one of the breakpoints $\{\hat{x}_i\}$.

It can be shown that if we wish to choose $\hat{a}$ such that

$$\sum_{i=1}^{k} (f(x_i) - y_i)^2 \tag{27}$$

is minimized, we have the problem of minimizing the euclidean length of the vector

$$A\hat{\mathbf{a}} - \mathbf{y}, \qquad \mathbf{y} = (y_1, \cdots, y_k)^T \tag{28}$$

subject to

$$E\mathbf{x} = 0 \tag{29}$$

and

$$C\mathbf{x} \geq 0 \tag{30}$$

The forms and dimensions of the matrices A, E, and C are straightforward, so we will not elaborate on this further, except to say that E and C reflect the continuity, convexity, and monotonicity requirements for f.

Computer-generated plotted curves[5] of an example which occurs in a JPL problem involving core flow predictions (Ref. 4) are shown in Figs. 7 to 9.

Figure 7 displays computer-generated graphs for f, f', f'' with the requirement that f is decreasing on $[-1,1]$, convex downward on $[-1,\hat{x}_m]$ and convex upward on $[\hat{x}_m,1]$.

Figure 8 displays graphs of the same derivatives with the additional requirement that f' be convex downward on $[-1,\hat{x}_m]$ and $[\hat{x}_m,1]$.

Figure 9 displays graphs of the first two derivatives with a less restrictive additional requirement that f' be convex downward on $[-1,\hat{x}_m]$ and convex upward on $[\hat{x}_m,1]$.

References

1. Rosen, J. B., "The Gradient Projection Method for Nonlinear Programming, Parts I and II," *J. SIAM*, Vol. 8, 1960, pp. 181–217; and Vol. 9, 1961, pp. 514–532.
2. Saaty, T. L., and Bram, J., *Nonlinear Mathematics*, McGraw-Hill, New York, 1964.
3. Hanson, R. J., and Lawson, C. L., "Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems, Part I," to be published in *Math. of Comp.* (in process).
4. Back, L. H., Witte, A. B., *Prediction of Heat Transfer From Laminar Boundary Layers, With Emphasis on Large Free-Stream Velocity Gradients and Highly Cooled Walls*, Technical Report 32-728. Jet Propulsion Laboratory, Pasadena, Calif., June 1, 1965.

[5]Computations performed by P. Breckheimer, JPL Computation and Analysis Sect.

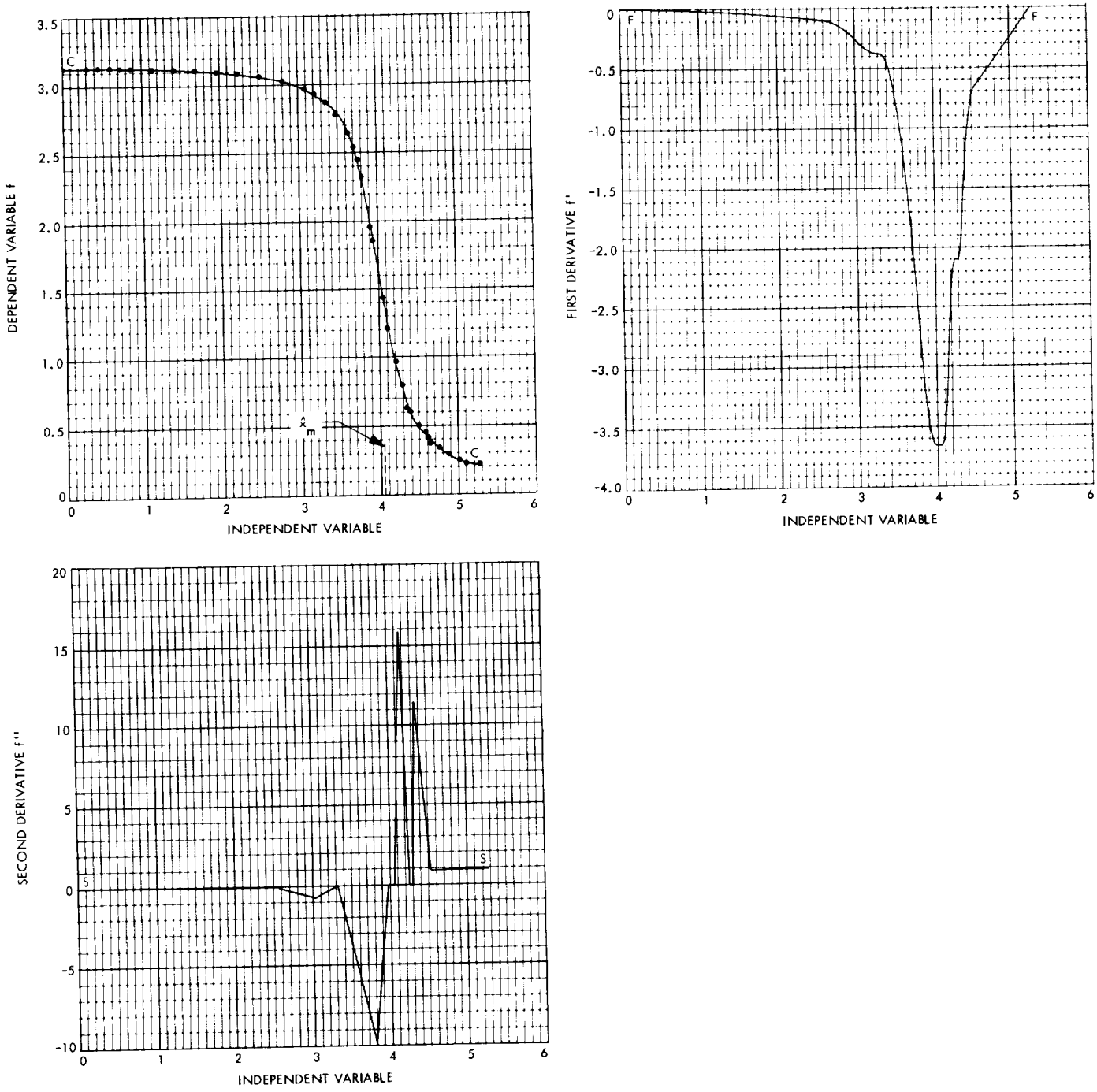


Fig. 7. Graphs for f, f', and f'' where f decreases on $[-1, 1]$, is convex downward on $[-1, \hat{x}_m]$, and convex upward on $[\hat{x}_m, 1]$

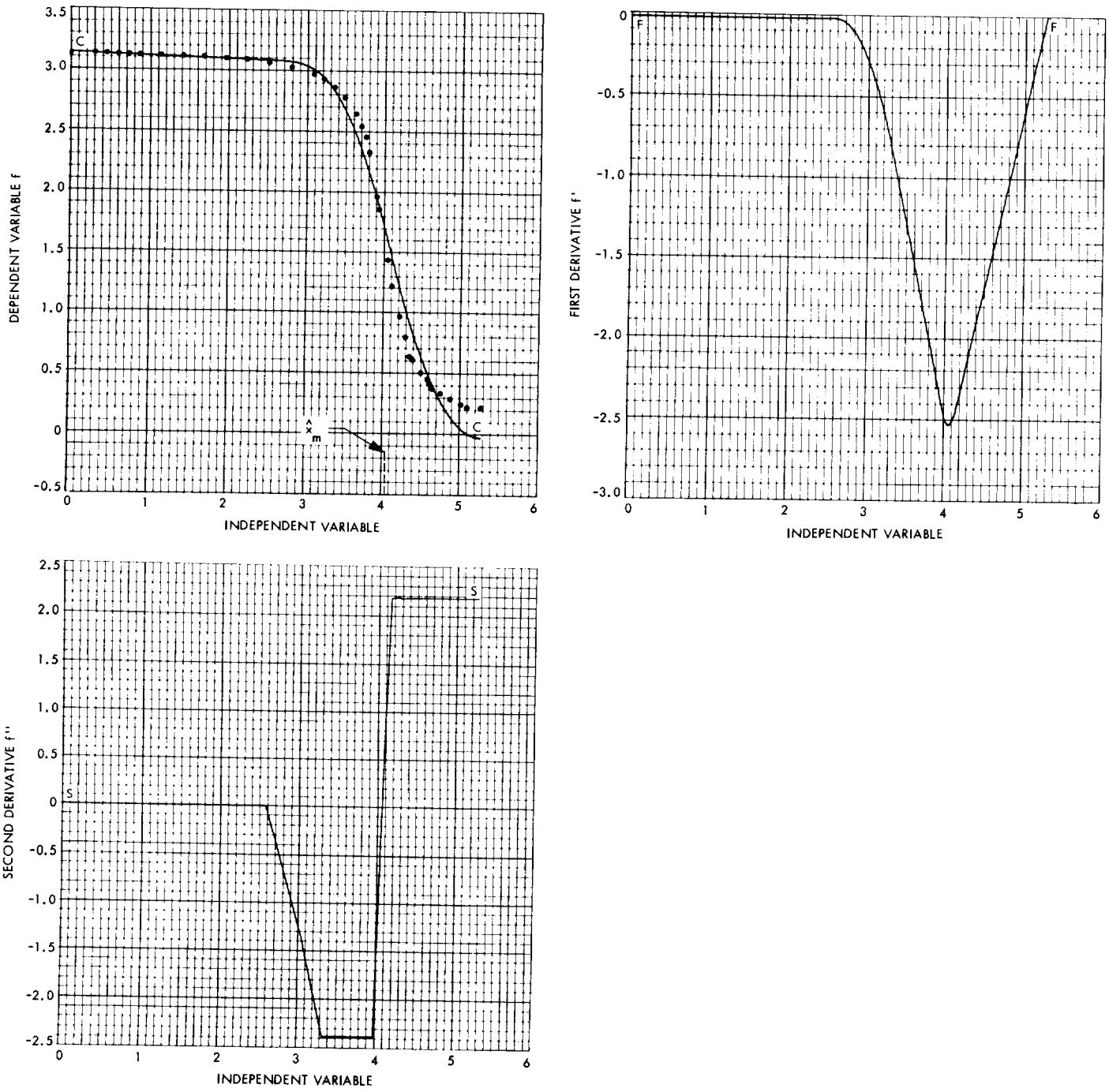


Fig. 8. Graphs for *f*, *f'*, and *f''* where *f'* must, in addition, be convex downward on $[-1, \hat{x}_m]$ and $[\hat{x}_m, 1]$

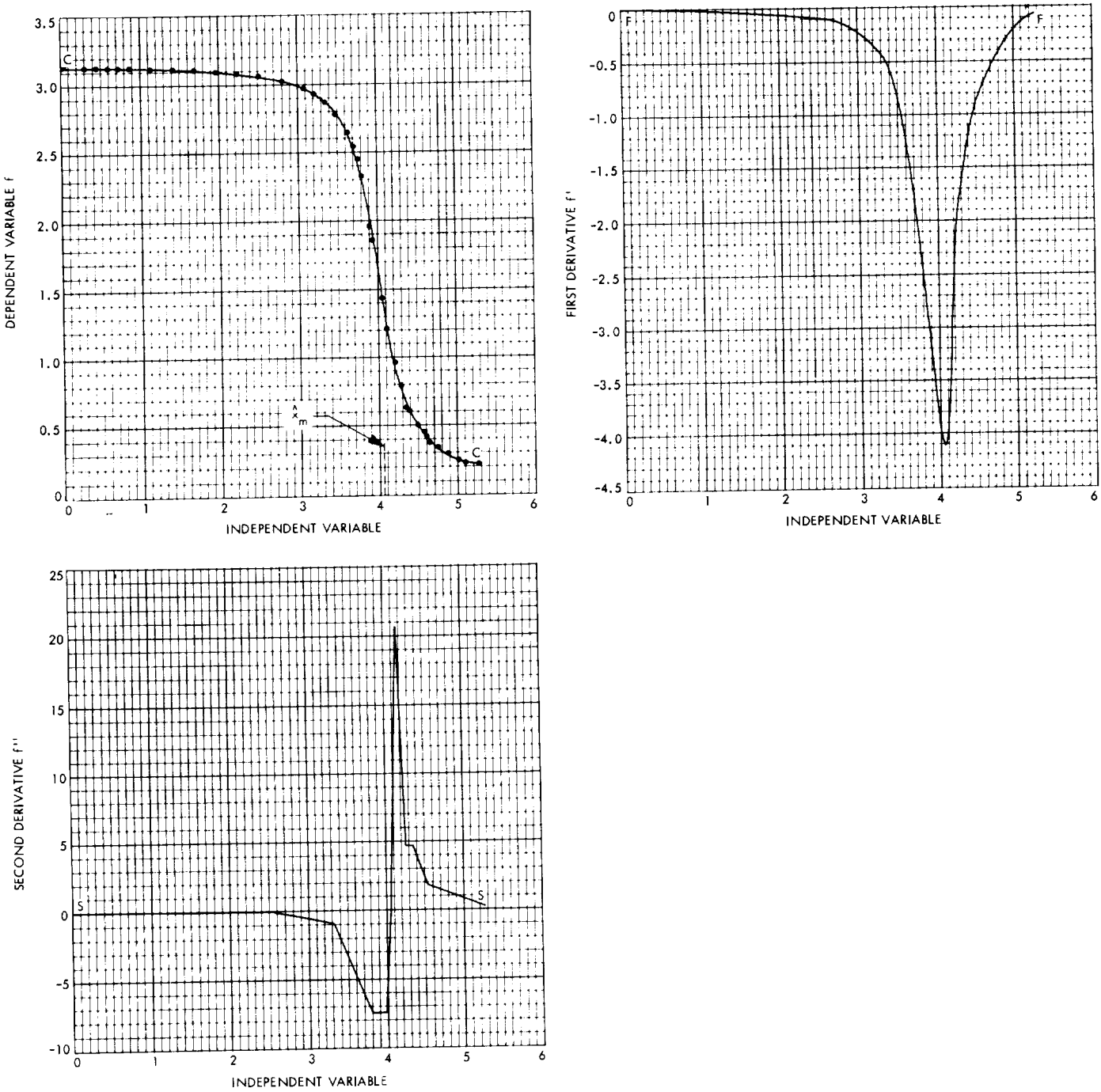


Fig. 9. Graphs for f, f', and f'' where f' must be convex downward on $[-1, \hat{x}_m]$ and convex upward on $[\hat{x}_m, 1]$

III. System Design and Integration

PROJECT ENGINEERING DIVISION

A. Entry and Landing Capsule System, *E. K. Casani*

1. Introduction

The objectives of the capsule system advanced development (CSAD) program were discussed in SPS 37-48, Vol. III, pp. 45–47. The design of a functioning engineering model, called the CSAD feasibility model, was described in SPS 37-49, Vol. III, pp. 67–76. In this article, the capsule system operational sequence is given, and the functional and environmental tests performed on the feasibility model are described.

The CSAD program has demonstrated that it is technically feasible to conduct a Mars entry and landing mission in the immediate future, and that the required technologies do exist. It has also shown that the type of equipment required for a high-g impact landing can be developed and the equipment required throughout the system can survive the heat cycle.

2. Operational Sequence

The capsule is mounted on a modified *Mariner* Mars 1969 spacecraft which is launched by an *Atlas/Centaur* vehicle into a trajectory which flies the spacecraft a predetermined distance by Mars. Ten days before planetary encounter, the capsule is separated from the spacecraft. After separation, a rocket motor is ignited to provide the required velocity increment. This maneuver deflects the capsule away from the spacecraft and puts it on an impact trajectory with the planet. During this separation and deflection phase, the capsule transmits engineering data to the spacecraft, which relays the data to earth. The last event in this phase occurs 15 min after separation, and then the capsule is turned off for the 10-day cruise period. This near-planet geometry is shown in Fig. 1.

At 15 min before entry into the planetary atmosphere, the capsule is turned on by an onboard timer and begins to transmit science calibration and engineering data to the spacecraft. As it enters the atmosphere, the capsule begins to simultaneously store and transmit both engineering and science data. Three-axis acceleration measurements are made on the capsule to determine the atmospheric density profile. A radiometer located in the nose of the capsule measures the intensity of the shock layer in front of the capsule at several different wavelengths to make gross compositional analyses of the atmosphere. Aerodynamic braking reduces the capsule speed to low supersonic levels; then a port in the front is opened

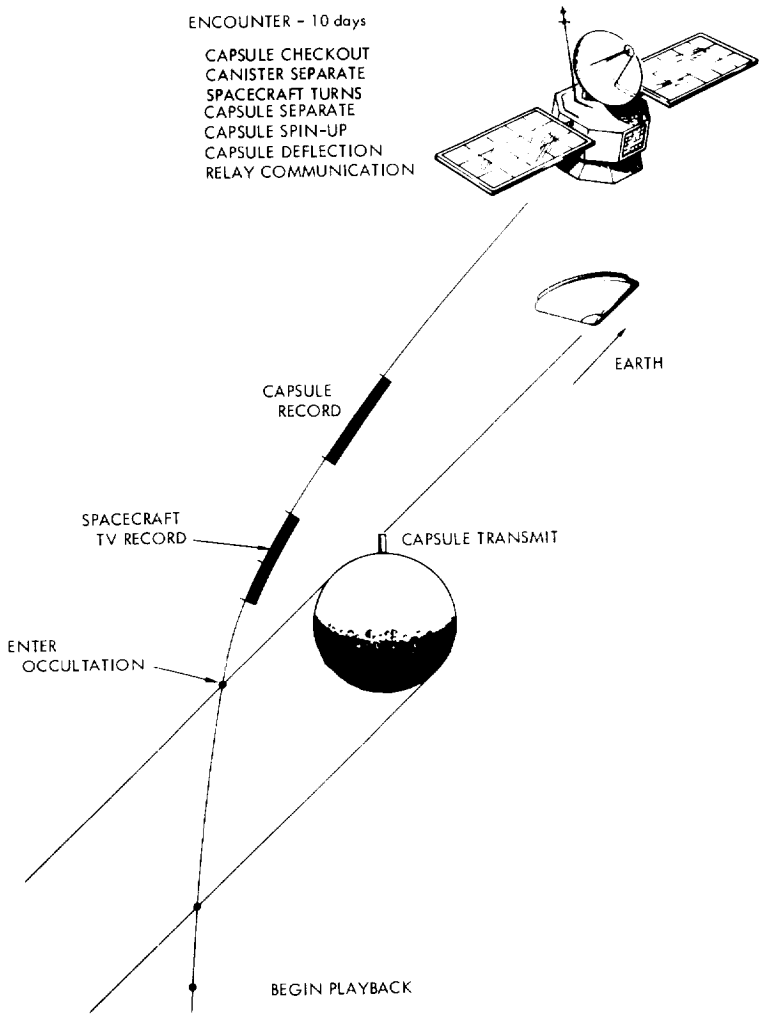


Fig. 1. Near-planet geometry

to take a sample of the atmosphere into the capsule for analysis by the mass spectrometer. Probes are extended through the heat shield to measure the low-altitude temperature and pressure profiles. (See Fig. 2, entry profile.)

As the capsule speed approaches transonic, a small lander is extracted from the capsule and descends to the surface of the planet on a parachute (Fig. 3). The capsule continues to transmit real-time science and engineering data to the spacecraft until impact, as well as continuing to transmit the stored data.

Upon landing, the parachute is released from the lander, and 30 s later, its radio begins to transmit directly to earth. Two and a half minutes after landing, an instrument boom is deployed from the lander to measure the surface wind-speed. For 20 min, the lander transmits engineering data along with wind, pressure, temperature, and water vapor measurements, and an atmospheric compositional analysis. At the end of this period, the lander radio is turned off for the Martian night. The next day, when the earth is again in view, engineering and scientific measurements which were made and stored during the Martian night are transmitted to earth.

3. Functional and Environmental Tests

A series of functional tests was performed on the capsule (Fig. 4). These tests were designed to operate the capsule in its different functional modes, including issuance of primary and backup commands, in different power system configurations, and in different radio and data handling modes. These tests demonstrated the functional adequacy of the capsule equipment prior to environmental testing.

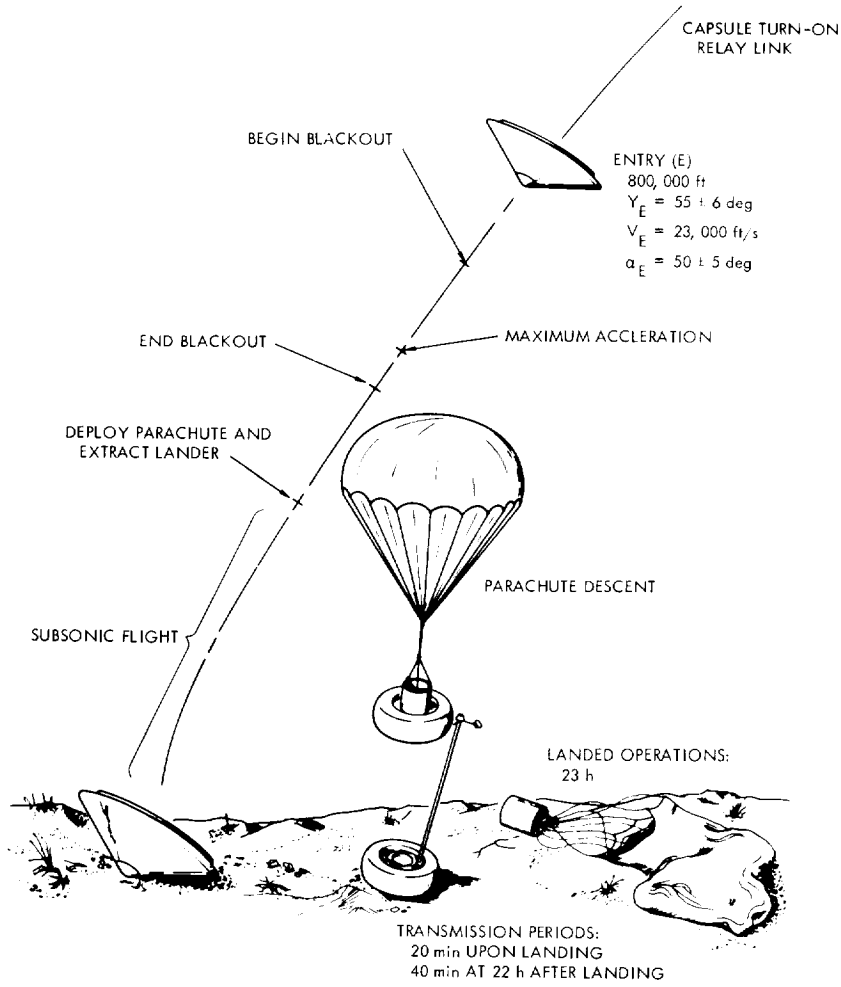


Fig. 2. Entry profile

The environmental test program consisted of subjecting the capsule to the two most difficult environments required for this type of mission: sterilization and impact.

If man is to look for life on Mars, it is important not to contaminate the planet with life from earth. To avoid this possibility, it is necessary to sterilize any capsule which would enter the planet's atmosphere. Today, the most acceptable sterilization technique is to heat the assembled capsule system to a temperature of 257°F in a sterilization chamber. When the capsule is in the chamber, it is enclosed in a sterilization canister so that after it is removed from the chamber and exposed to the earth's atmosphere, only the external surface of the canister becomes recontaminated. The capsule remains in this canister during its flight through space and up until several days before it enters the planet's atmosphere; at this point, the canister is opened and the sterile capsule is separated from the spacecraft.

The sterilization environmental test was conducted by placing the capsule in its sterilization canister into the chamber for 35 h; the chamber was filled with gaseous nitrogen, and its temperature was raised to 257°F.

After completion of the sterilization test, a capsule system functional test was conducted. The lander was then removed from the entry aeroshell and taken to the Goldstone Deep Space Communication Complex for a high-velocity impact test. The lander was carried by a helicopter to an altitude of 250 ft and allowed to freefall to the desert floor. Tests were conducted on the "hardpan" bed of a dry lake simulating the Mars landing conditions and also on an asphalt roadway to simulate the design conditions. In both of these tests, the impact velocity of the 63½-lb lander was 80 mph before it struck the ground. On a mission to Mars where the atmosphere is about 200 times thinner, the landing craft would be slowed to the same velocity by a parachute.

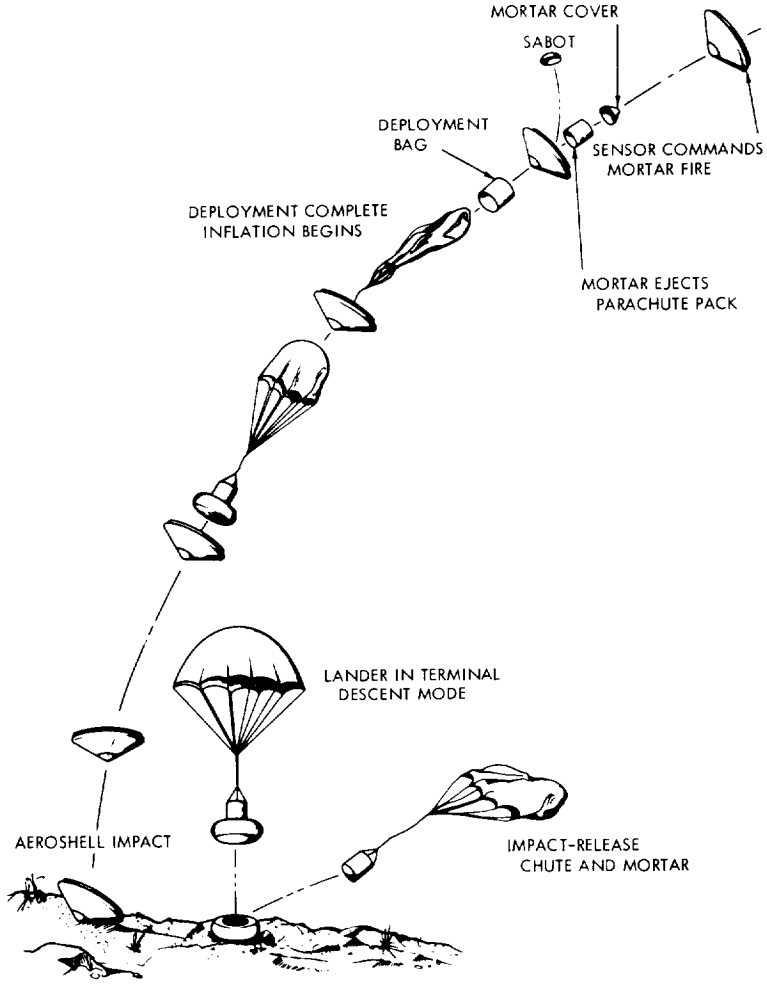


Fig. 3. Parachute sequence

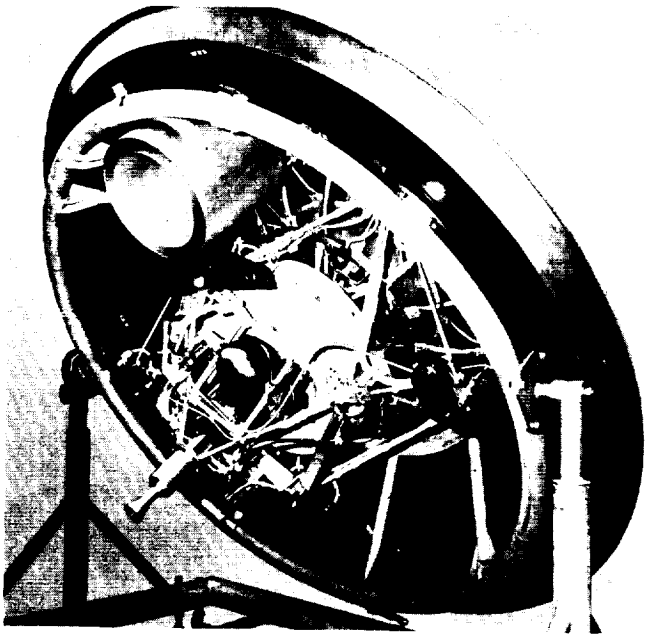


Fig. 4. Entry and lander capsule (separation configuration)

The radio transmitter turned on automatically 30 s after the lander struck the surface and operated for a scheduled 20 min. At 2½ min after impact, a tiny anemometer—a wind velocity detector—deployed at the end of a 4-ft telescoping instrument boom. A wind instrument is one of the prime experiments under consideration for initial planetary landing missions. Other experiments include investigation of surface pressure, temperature, water vapor, and low-mass atmospheric constituents.

On a mission to Mars, as in the Goldstone test, the Mars landing craft would be cushioned by balsa wood to absorb landing shock. In the test conducted on the dry lake bed, the impact energy was mostly absorbed by the lake bed, whereas in the test conducted on the asphalt, the energy was absorbed by crushing and shattering the balsa wood.

Following a mission profile identical to projected Mars surface operations, the lander's radio turned on again 22 h after the initial transmission. Signals were received for another 40 min to conclude the test.

IV. Flight Computers and Sequencers

GUIDANCE AND CONTROL DIVISION

A. STAR Computer Assembler and Loader,

J. A. Rohr

1. Introduction

The JPL STAR computer currently being developed is a Self-Testing-And-Repairing computer. The organization of this machine is described in Refs. 1 and 2. The use of the computer in unmanned interplanetary travel for guidance and control computation, as well as onboard processing of scientific data, is described in SPS 37-46, Vol. IV, pp. 57–62. An overview of the software system for the STAR computer is presented in SPS 37-50, Vol. III, pp. 75–77. The assembler and loader are described in this article.

The first two programs of the STAR computer software system, SCAP (STAR Computer Assembly Program, i.e., assembler) and LOAD (STAR computer loader), are used to prepare programs to be run on the STAR computer. At the present time SCAP is complete except for macro facilities and certain pseudo-operations pertaining to input–output operations and subroutine linkage. LOAD is complete at this time. The primary design objective of these programs is completeness and versatility for the STAR computer programmer. A secondary objective is minimization of programming effort without any sacrifice of the primary objective.

2. Assembler

SCAP is designed both to provide all the usual features found in an assembler for a conventional computer, and to facilitate later modification for automatic insertion and checking of rollback points as described in *Subsection 2-g*. The most unique feature of SCAP is the COMPILE pseudo-operation which automatically compiles certain types of arithmetic expressions. Macro facilities are included in the design of SCAP, but will not be implemented until a later date. To simplify programming of SCAP, the symbol and literal tables use a fixed-size format. To make the most efficient utilization of the remaining memory space, the reference tables use a linked-list format. The symbol and literal tables have a binary tree format for Pass 1 and a binary list format for Pass 2. The reference tables appear in Pass 2 only. The external symbol linkage used in SCAP and the loader is basically a linking-loader type (see *Subsection 3*).

Instructions assembled by SCAP are written in SCAL (STAR computer assembly language). A SCAL statement for a machine operation consists of a label field, an operation field, an address field, a tag field (if required by the operation), a comments field, and an identification field. The tag field is used only for instructions which may have an index register specified. A SCAL statement

for a pseudo-operation consists of a label field, an operation field, a variable field consisting of one or more subfields, a comments field, and an identification field. The particular format depends on the pseudo-operation.

a. Organization. SCAP is organized as a two-pass assembler. The input to SCAP is a card deck of SCAL statements. Prior to reading the first card, files and variables are initialized and default options are set. During Pass 1, the symbol and literal tables are constructed, header card information is accumulated, and macros and COMPILE pseudo-operations are expanded. The intermediate processing between Pass 1 and Pass 2 produces the title and header cards for the deck and rearranges the symbol and literal tables. During Pass 2, the address and tag for each instruction are evaluated and the instruction is constructed. Finally, after Pass 2, the literal table, symbols and symbol reference table, all undefined symbols and their references, all multiply-defined symbols, an error legend, and a summary are printed.

b. Pass 1. The input to Pass 1 is supplied by a subroutine which provides card images from the input deck, the macro expander, or the COMPILE expander. The operation specified for the card is used as the key for searching the operation table. If the operation is a macro specification or a macro call, control passes to the macro processing subroutines. If the operation is a pseudo-operation, control passes to the appropriate subroutine for Pass 1 processing. If the operation is a machine operation, the following action is taken. First, the label field is checked. If a valid symbol appears in the label field, it is entered into the symbol table. Next the address field is checked for a literal. If a legal literal is used in the address field, it is entered into the literal table if it is not already there. Next the address field is checked for an external symbol. If the address consists of an external symbol only, this reference is noted in the symbol table. Finally, the input card and two control words are written onto an intermediate file.

c. Interlude. The intermediate processing between Pass 1 and Pass 2 produces title and header cards and rearranges the symbol and literal tables. The first card of the output deck is a title card containing the deck name, date of assembly, initial instruction address, program length, COMMON length, number of entry points, and number of external symbols for the deck. Header cards follow the title card. The header cards specify each entry point and its location, and each external symbol and the location of the base of its internal linkage chain. The symbol and literal tables are transformed from the binary tree format used for Pass 1 into the binary list format used for Pass 2.

d. Pass 2. The second-pass processing is primarily devoted to construction of machine instructions in an intermediate language used as input to the loader. An instruction in the intermediate language consists of an operation code that includes the tag specification, an address, and three bits that specify the relocation attributes of the address. One bit is used to specify program relocation, one to specify COMMON relocation, and one to specify the position of the address field in the word. When a card is read from the intermediate file, the operation code information passed from Pass 1 is examined to determine whether a machine operation or pseudo-operation is specified. If a pseudo-operation is specified, control passes to the appropriate subroutine for Pass 2 processing. If the operation is a machine operation, the address is evaluated, the tag (if present) is evaluated, and this information, and the operation code information passed from Pass 1, is used to construct the instruction. The instruction is then printed and/or punched according to the listing and punching options in effect at the time.

e. Postlude. The post-processing phase of SCAP provides the programmer with useful information accumulated during the assembly. First, the literals generated by the assembly are listed. Next, a symbol and symbol reference table are printed. This table lists each symbol, its value, the card number of its definition, and the number of each card on which the symbol appears in the address, tag, or variable field. Next, all undefined symbols, and all references to them, are printed. Then, all multiply-defined symbols and the card numbers of the definitions are printed. Finally, an error legend and summary are printed. The summary lists the input and output card count, output page count, and error counts.

f. Pseudo-operations. Most conventional pseudo-operations are included in the set of SCAP pseudo-operations. Among these are list- and punch-control pseudo-operations for selective control of the printed and punched output; storage reservation pseudo-operations, including a COMMON feature similar to that of Fortran; and data-loading pseudo-operations. The generation of data by data-loading pseudo-operations includes automatic encoding of each data word as required by the design of the STAR computer for error checking. Decimal, hexadecimal, and binary constants may be generated as well as address-type constants. An EQU (equate)

pseudo-operation is included to set a symbol equal to an expression, and a SET pseudo-operation is included which is the same as EQU except that redefinition is allowed. The final value of a SET symbol in Pass 1 is used as its value throughout Pass 2. Two pseudo-operations, ENTRY and EXTER (external) are used for linkage specification. For each assembly, each entry point and each external symbol must be explicitly specified. An entry point is a symbol whose location is available to all decks of a load; an external symbol is one whose value is defined by an entry point in another deck. The COMPILE pseudo-operation implements compilation of ordinary arithmetic expressions. Any syntactically legal expression using =, +, −, *, /, and ( ) operators is allowed. The programmer must define elsewhere all variables except working storage. The COMPILE feature greatly increases the versatility of SCAP.

g. Rollback. The self-repairing features of the STAR computer are implemented primarily in the hardware of the machine as part of its normal operation. When an error occurs, the computer automatically returns the program to a prespecified rollback point and tries to do the calculation again. If the error does not repeat, the trouble is assumed to be transient in nature, and computation continues. If the error does repeat, the faulty module is replaced and the program is again resumed at the rollback point. Initially, the programmer will be completely responsible for the specification of these rollback points. One research area of the STAR computer project is an investigation of automatic checking and insertion of rollback points by SCAP in a STAR computer program.

3. Loader

LOAD is a relatively straightforward program which assigns absolute values to relocatable addresses and performs external symbol linkage. The deck-name table uses a fixed-length format. In order to get maximum utilization of the available core space for reference tables, the symbol table, and all reference tables, use linked-list formats.

Each symbol in a deck referenced by other decks of the LOAD is specified on a header card along with its location in the deck. The symbol table is constructed using all entry points defined in the LOAD. For each external symbol used in each deck, a header card specifies the symbol and the location of the base of a linkage chain. This chain links every location in the deck that uses the symbol. LOAD uses this chain to replace each address in the chain by the actual value of the symbol as defined by an entry point in one of the decks of the LOAD.

a. Organization. LOAD is basically organized in two passes. The first pass actually loads each instruction into the memory, adding program relocation to the address if specified; generates the symbol table from entry-point specifications; and prepares an intermediate file of external-symbol specifications. The second pass links external symbols and relocates all COMMON addresses. Finally, a postlude prints undefined and multiply-defined symbols, a cross-reference table, and a core map.

b. Pass 1. The input to LOAD consists of one or more decks assembled by SCAP and output in an intermediate language. Each deck begins with a title card followed by header cards and program text. The title card information is entered into the deck name table. The entry points on the card are entered in the symbol table. The external symbols are recorded on an intermediate file so that the input deck does not have to be read more than once. Each word of the program is entered into the memory as it is read. If program relocation is specified for the address, it is added at this time. if COMMON relocation is specified, a flag is set.

c. Pass 2. During Pass 2, each external symbol address is set and COMMON relocation is added where specified. Each external symbol recorded on the intermediate file is read and the base of its linkage chain established. The value of the symbol is obtained from the symbol table. Beginning at the base and following the linkage chain, the value is set in each address field in the chain. After all external symbol values have been set, the entire memory is scanned for words having the COMMON relocation flag set. On all such words, the COMMON relocation amount is added to the address. The loading is then complete.

d. Postlude. The post-processing phase of LOAD provides the programmer with useful information accumulated while loading. A table of all undefined symbols, and every location in which they are used, is printed so that the address may be supplied from the STAR computer console. A table of all multiply-defined symbols, and the decks in which they are defined, is printed. An option allows the printing of a cross-reference table which lists, for each deck, the name, base, and length of the deck, each entry point in the deck with the names of the decks referring to it, and each external symbol with its value and the deck in which it is defined. A final option is a core map that gives the location of each deck and symbol of the LOAD.

References

1. Avizienis, A., Rennels, D. A., and Rohr, J. A., "Application of Concurrent Diagnosis and Replacement in a Self-Repairing Computer," Presented at IEEE International Convention in New York, Mar. 18–21, 1968.

2. Avizienis, A., "The Design of Fault Tolerant Computers," in *AFIPS Conference Proceedings*, Vol. 31, (Fall Joint Computer Conference), pp. 733–743, Thompson Books, Washington, D.C., 1967.

V. Spacecraft Power

GUIDANCE AND CONTROL DIVISION

A. Thermionic Converter Technology, *O. S. Merrill*

1. Introduction

Previous SPS articles[1] have reported a contracted effort to develop planar thermionic converters having optimized performance in close agreement with the performance predicted from the data obtained on a variable parameter test vehicle. The reason for developing converters to operate at a specified voltage and power output is that these factors must be known with confidence in advance in order to design thermionic spacecraft power systems with predictable performance characteristics. This article gives a description of the test vehicle, the converters, and the correlation of performance between them.

2. Variable Parameter Test Vehicle

Figure 1 shows a cross section of a thermionic research test vehicle depicting the critical elements of the structure. There are three operational features of this vehicle which are noteworthy:

(1) The test vehicle is so designed that the emitter temperature can be measured with an NBS-calibrated[2] micro-optical pyrometer sighted (direct line of sight) into a 10:1 blackbody hole positioned in the side of the emitter. In this manner, no extraneous radiation from the electron bombardment heater can introduce error into the emitter temperature measurement.

(2) The device proper is thermally isolated from the cesium reservoir by the thin-walled cesium tubulation. It is, therefore, possible to maintain the device at a temperature which is 200°C or more higher than the cesium reservoir, which insures that the cesium reservoir temperature uniquely establishes the cesium atom arrival rate at the emitter, collector, and guard ring surfaces.

(3) The interelectrode spacing of the test vehicle is capable of adjustment over the range from 0.1 to

[1]SPS 37-39, Vol. IV, pp. 15–19; SPS 37-50, Vol. III, pp. 82–92; SPS 37-51, Vol. III, pp. 41–44.

[2]NBS — National Bureau of Standards.

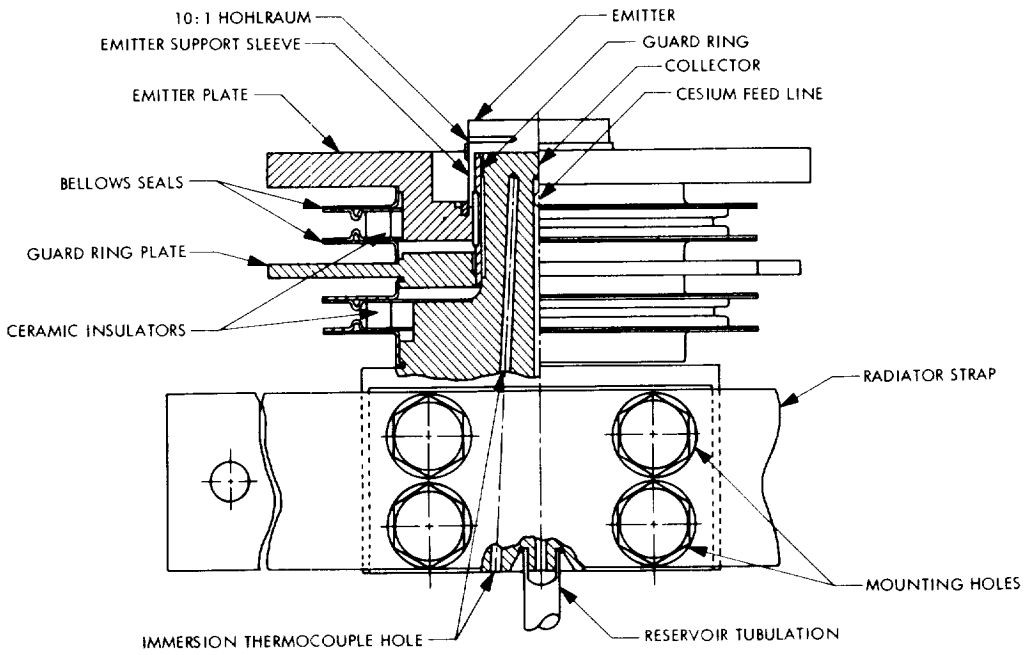


Fig. 1. Variable parameter test vehicle

15 mils and can be established to within ±0.05 mil at any spacing by a precision drive and indicator mechanism. The interelectrode spacing is adjustable at three equally spaced points to achieve parallelism. To maintain parallelism during operation, the individual adjustments can be locked together and driven from one central location.

The test vehicle is fabricated from high-temperature ceramics and refractory metals such as niobium, molybdenum, rhenium, and tantalum. The subassemblies (Fig. 2) are prefabricated by high-temperature brazing with vanadium or titanium and assembled by electron-beam welding. The electrodes are preprocessed at temperatures 400°C higher than intended operation in order to stabilize the surface structure. As a final step, the device is exhausted to a terminal pressure of 5×10^{-8} torr at bake-out temperature and then loaded with triple-distilled 99.98% purity cesium. The completed vehicle is mounted in an ion-pumped vacuum chamber and instrumented for test.

A polycrystalline rhenium emitter and collector were selected for initial performance evaluation. During the test program, various types of data were obtained such as saturated electron emission, converter optimization, and voltage output (power output) versus interelectrode spacing. These latter data were obtained by instrumenting the test vehicle to measure the voltage or power output variation as a function of spacing between the electrodes at a constant current level with a constant emitter, cesium reservoir, and collector temperature. In turn, these non-varying conditions insure a constant sheath thickness at the emitter and collector, a constant plasma density, and a constant plasma electron temperature. Therefore, the voltage output variation with spacing may be related to a power output profile of the vehicle interelectrode space. Moreover, the spacing is easily controllable to within ±0.05 mil since all the test vehicle thermal expansions are fixed by the constant heat load through the electrode structures.

3. Fixed-Spacing Converters

The converter design (Fig. 3) purposely retained as many of the common features of the test vehicle as possible.

Figure 4 illustrates the similarity of fabrication and external detail of emitter structure and blackbody hole geometry. Similar emitters, hohlraums, metal-ceramic seals, divided current-lead straps, and thermally isolated cesium reservoirs are some of the more readily observable common features. The philosophy is that as many design and operational aspects of the test vehicle as possible should also be incorporated into the converter, for only

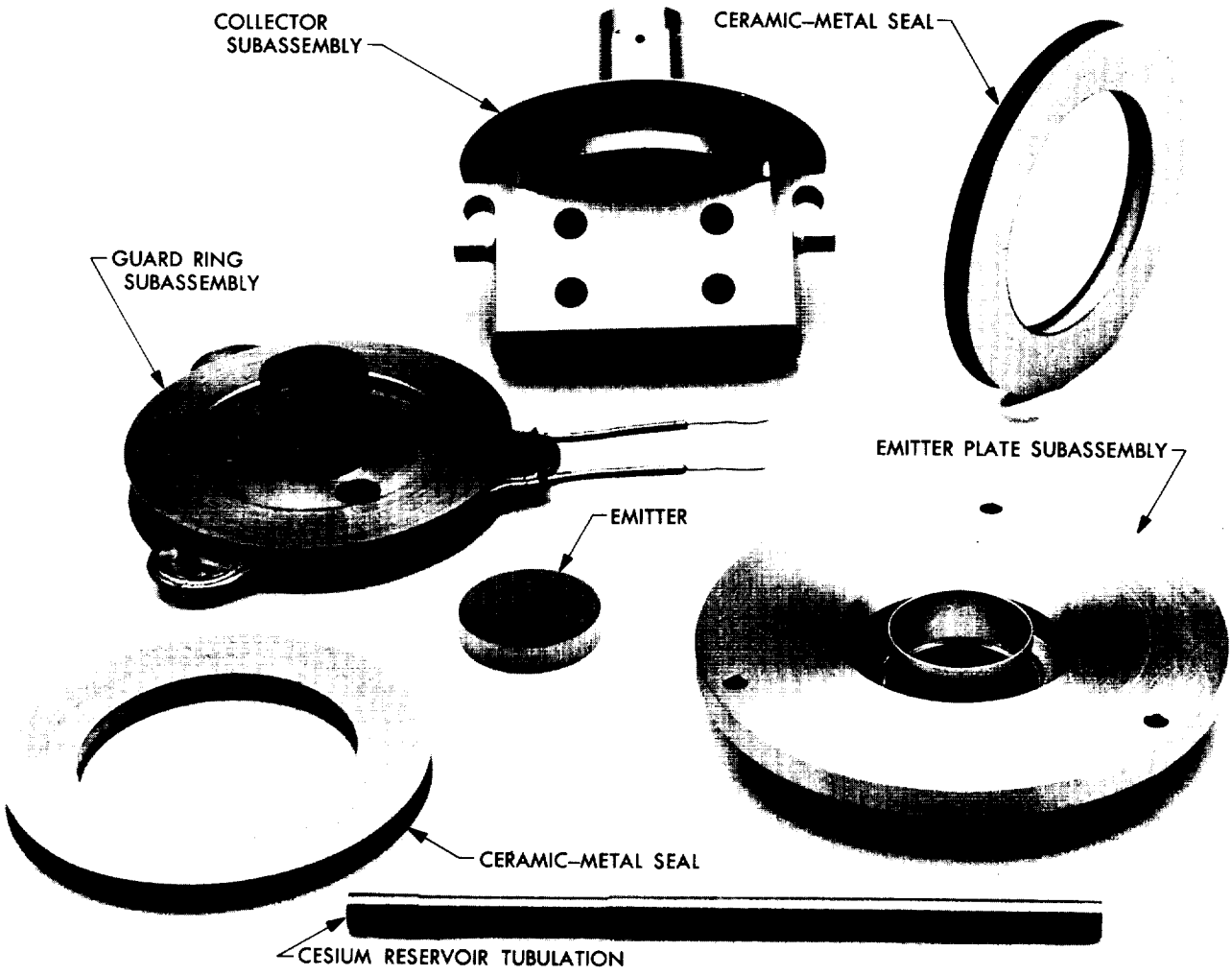


Fig. 2. Variable parameter test vehicle subassemblies and components

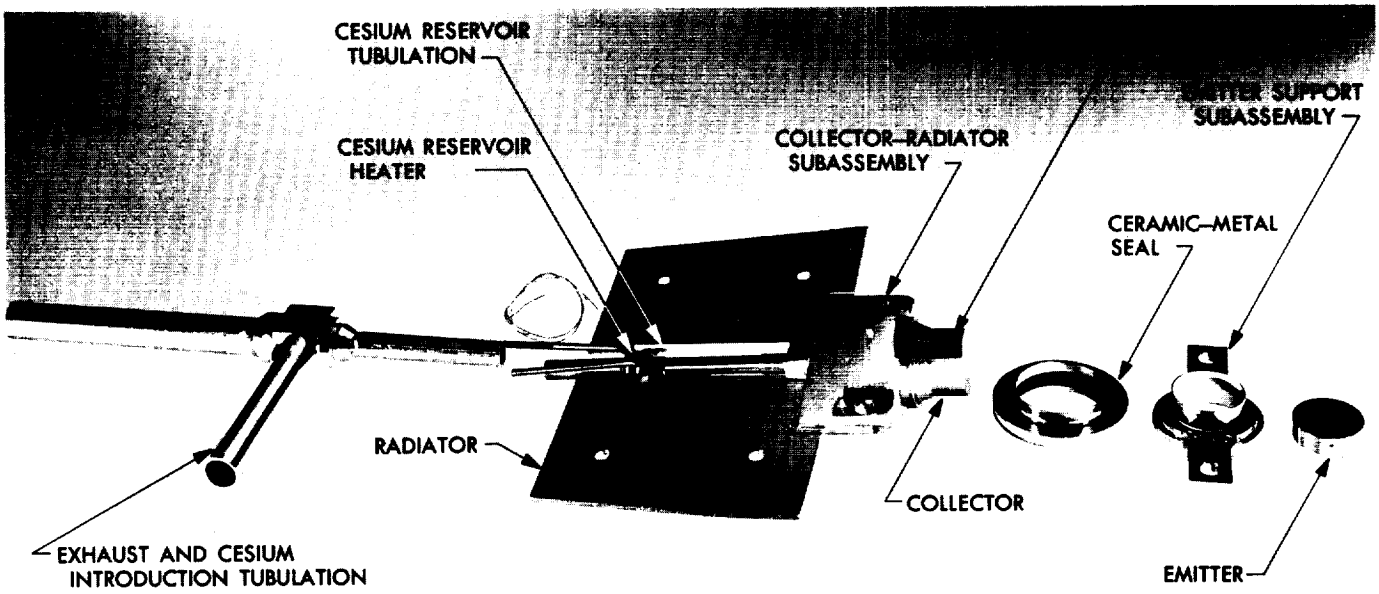


Fig. 3. Converter SN-101 subassemblies and components

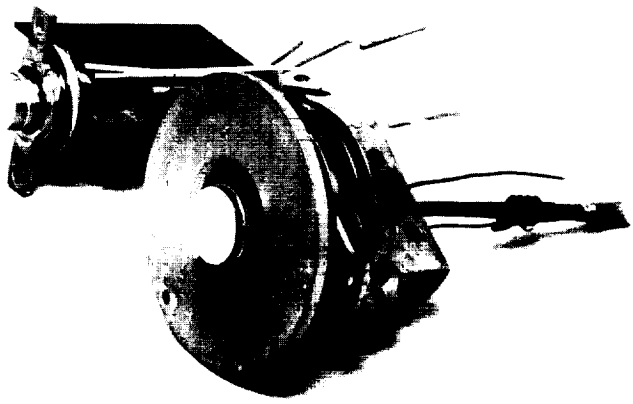


Fig. 4. Test vehicle (foreground) and converter SN-101

in this manner is a one-to-one comparison of performance possible. Particular attention was accorded to minimizing the sidewall emission of the converter (Table 1). The test vehicle is guard-ringed to prevent this. However, hardware design precludes the use of three electrode elements and two ceramic insulators, as in the test vehicle.

The fabrication of the SN-101 series converters represents advances in thermionics converter fabrication—prefabrication of the ceramic-metal seals, their pretest, and their final assembly into the converter configuration by electron-beam welding. Converters built previously were dependent upon the ceramic–metal sealing operation as the final assembly step. Two disadvantages resulted: (1) a seal failure generally rendered one subassembly, and perhaps two, useless with minimal chance of recovery; and (2) the vapor pressure of seal brazing material was often high enough at this melting point that the converter interior, including the electrodes, was coated with several layers of contaminant materials.

The collector–radiator subassembly is comprised of a molybdenum collector barrel, OFHC[3] copper radiator fins, a niobium welding ring, a rhenium collector shim, and a tantalum reservoir. The rhenium shim is vanadium-brazed to the collector barrel since the resultant bond is made without the formation of a brittle intermetallic or low-melting-point eutectic alloy. The other components are titanium-brazed. The niobium welding ring is selected to provide a base material identical to that of the sealing flanges to avoid the apparent incompatibility of a solid solution of niobium and molybdenum formed during electron-beam welding.

The emitter support subassembly is comprised of a rhenium tube, a niobium transition ring, and tantalum lead straps. All parts are vanadium-brazed. The niobium transition ring contains a 0.060-in.-thick lip section to which the upper flange of the prefabricated seal is electron-beam welded.

[3]OFHC = oxygen-free high conductivity.

Table 1. Comparison of converter and variable parameter test vehicle design features

Design or operational particular	Converter	Variable parameter test vehicle
Measurement of emitter temperature	Blackbody (8:1 to 10:1) hohlraum; directly observed in line of sight; no extraneous radiation from heater filament	Blackbody (10:1) hohlraum; directly observed in line of sight; no extraneous radiation from heater filament
Method of heating emitter	Indirect heating by counterwound, pancake electron bombardment heater	Indirect heating by counterwound, pancake electron bombardment heater
Power-producing region	Collector area and corresponding emitter area only; sidewall emission minimized by heat choke directly below emitter and wide spacing between envelope and collector	Collector area and corresponding emitter area only by using guard ring
Cesium reservoir	Isolated from converter structure to establish a unique reservoir temperature	Isolated from test vehicle structure to establish a unique reservoir temperature
Method of attaching emitter	Solid rhenium emitter electron beam welded into rhenium tubing	Solid rhenium emitter electron beam welded into rhenium tubing
Interelectrode spacing	Fixed spacing by positive method; ±0.0003-in. accuracy	Variable spacing with drive mechanism; ±0.0005-in. accuracy

The final assembly operation is the electron-beam welding of the rhenium emitter to the emitter support subassembly. The weld schedule follows previously established parameters for electron-beam welding rhenium, i.e., an electron-beam voltage of 150 kV at a current of 4.9 mA, with a part revolution speed of 40 rpm. Attaching the emitter last has several advantages:

(1) It allows for visual inspection of the concentricity of the envelope and collector.

(2) It permits a high-temperature emitter outgas and grain stabilization treatment independent of the converter structure. In particular, the outgassed emitter impurities will contaminate the collector surface and reduce the converter voltage output.

(3) It provides a means of setting the interelectrode spacing to a predetermined value by grinding or shaping the emitter before welding.

In the final converter assembly depicted in Fig. 3, only three electron-beam welds are necessary for assembly of the converter—two welds at the flanges and one at the emitter. The exhaust tubulation is joined to the diode assembly and high-purity cesium is distilled into the converter. The final step is the application of Rokide "C" ($\epsilon \simeq 0.78$) to the radiator fins to enhance the removal of waste heat.

4. Comparison of Test Vehicle Data With Converter Performance Data

a. Selection of design criteria. On the basis of the rhenium–rhenium test vehicle data shown in SPS 37-51, Vol. III, p. 42, Fig. 12, design criteria were selected for the fabrication of four additional fixed-spacing converters for emitter temperatures of 1327, 1427, 1527, and 1627°C. Other stipulated conditions were:

(1) Each converter was to have a polycrystalline rhenium emitter and collector, the emitter area being 2 cm^2 as defined by the inside diameter of the emitter sleeve.

(2) Each converter was to be so designed as to make sidewall emission negligible (defined as less than 10%) compared to that of the emitter.

(3) The radiator fins for each converter were to be sized for the current load and the emitter temperature to allow optimized performance while operating passively at the design point, i.e., no additional heat input to the collector–radiator assembly.

(4) Each converter was to have an independent cesium reservoir and a heater capable of heating the reservoir over the entire range of operating temperatures commensurate with the design criteria.

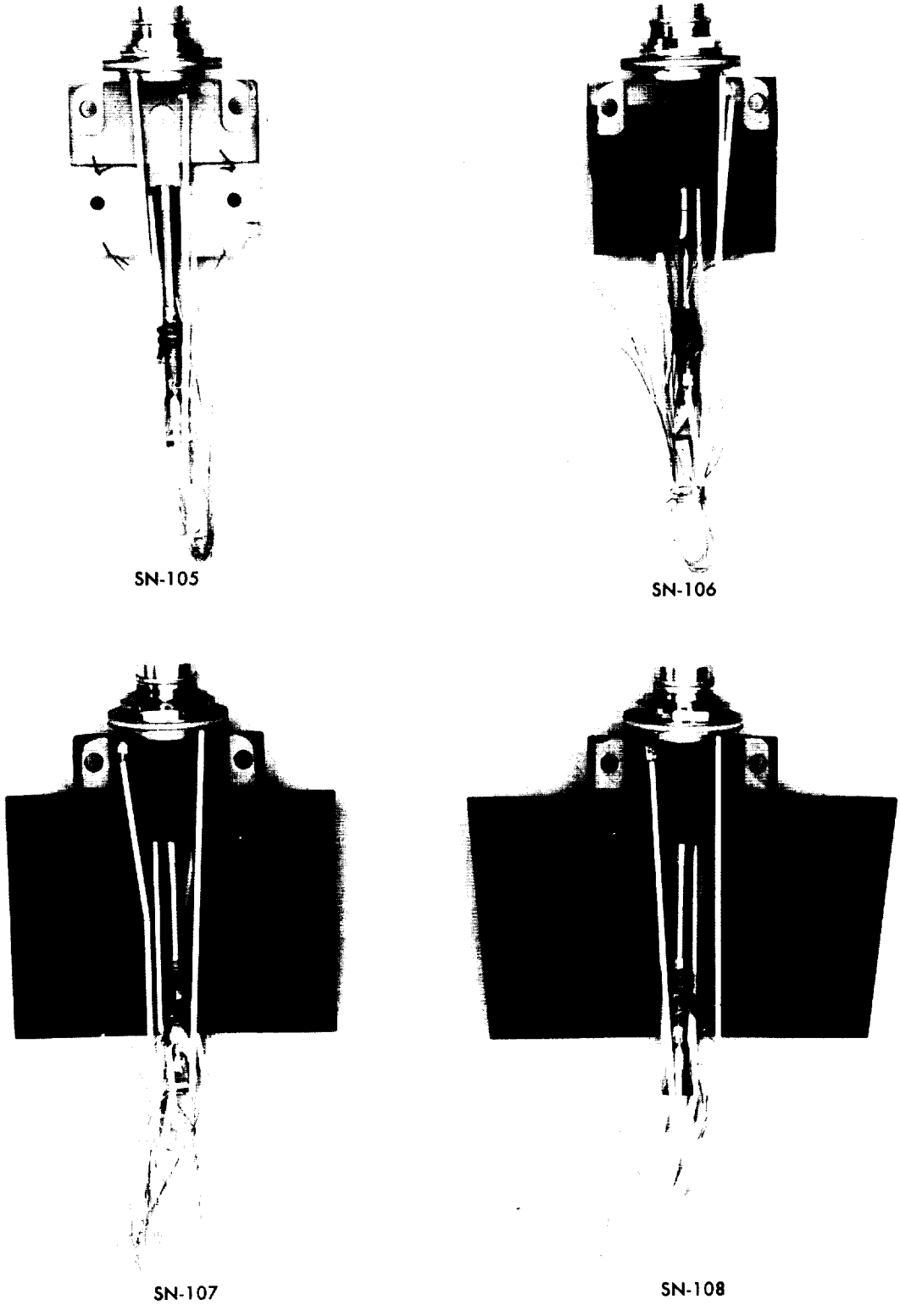


Fig. 5. Converters SN-105, -106, -107, and -108

Table 2. Comparison of converter performance with design criteria

Converter number	Emitter temperature, °C	Interelectrode spacing, mils	Output voltage, V		Output power density,[a] W/cm²		Efficiency observed,[b] %
			Design	Observed	Design	Observed	
SN-105	1327	12 ± 2	0.2	0.2	2.5	2.54	4.23
SN-106	1427	8 ± 1.5	0.3	0.3	5.25	5.59	5.85
SN-107	1527	5 ± 1	0.4	0.4	11.0	11.4	8.85
SN-108	1627	4 ± .5	0.55	0.52	15.0	15.65	9.30
SN-101	1735	3.2 ± .5	0.7	0.7	21.3	21.0	12.22

[a]On the basis of the 2-cm² emitter area. If the output power is desired on the basis of the 1.88-cm² collector area, multiply these values by the factor 1.064.
[b]Efficiency is defined as *output power/bombardment power*. It does not account for any losses or geometry factors and is therefore conservative.

(5) Other conditions specified instrumentation, methods of measuring temperatures, and fabrication and processing procedures, all of which were to be identical with the corresponding factors for converter SN-101.

b. Results. In Table 2, the performance of the four converters is tabulated along with that of converter SN-101, showing both the design criteria and the observed performance of the converters. A comparison of this data shows good agreement, which is significant both from a performance point of view and from the point of view of fabrication, processing procedures, and experimental technique.

The four converters, SN-105 through SN-108, are shown in Fig. 5. The relative sizes of the radiators are here dramatically illustrated. Note that for SN-105 and SN-106 the radiators are approximately the same size, 1⅞ × 2 in. overall, but the radiator for SN-105 is completely devoid of the high emissivity Rokide "C" coating. This was necessary in order for the converter to have sufficient radiator area to which one could attach thermal heating units for increasing the radiator temperature and thus the collector temperature for parametric testing. Such testing has not yet been done, nor has the extensive performance and parametric testing which is planned. These tests will explore, among other things, off-optimum performance caused by a variation in collector temperature, emitter temperature, and cesium reservoir temperature.

B. Low Saturation Drop Transistor, *A. I. Schloss*

1. Introduction

There are presently available efficient electrical sources of energy whose output voltages are in the 2- to 5-V range. Since this range of voltage is impractical for most useful loads, the voltage must be converted upward efficiently. Dc to dc conversion can be accomplished using transistors as switching elements to convert dc to ac. Efficient conversion requires that the switch have low internal resistance while *on*. Germanium transistors have this characteristic; however, allowable ambient temperatures are limited for germanium. A greater temperature margin is possible using silicon transistors; however, to date silicon has exhibited rather high internal resistance. It is the purpose of this effort to develop a low saturation drop silicon transistor (i.e., low resistance).

Effort to produce such a transistor has occurred in two stages. Phase I required a saturation drop of 0.2 V at 75 A. Phase I was completed. Phase II requires a saturation drop of 0.1 V at 75 A and the effort is continuing. Several approaches to achieve this goal are outlined in this article.

In phase I, ITT and Westinghouse Electric Corp. used two different basic approaches. ITT used a multichip approach and Westinghouse used a single large area star geometry approach. ITT would not bid on phase II and Westinghouse has continued in an effort to improve their techniques. A "spin off" of this effort has resulted in a higher voltage version of the transistor now offered commercially.

2. Thin Wafer Approach

The earlier (0.2-V) effort used a simultaneous diffusion process. It was decided to use this process in conjunction with a thin crystal. The thin crystal permits steeper concentration gradients, thereby increasing injection efficiency, a factor in low voltage drop. Planar technology is used in fabricating the device. Breakage during fabrication is a major problem with such an approach.

3. Two Controlled Collector Geometry Approaches

Current crowding at the emitter edge contributes to higher voltage drops. By controlling geometry such that the collector and emitter geometry are identical, it is possible to attain symmetry of gradients in the base under high injection, thus forcing equality of collector to base and emitter to base junction drops. Low saturation voltage drop is attained when these two drops are nearly equal. One method to achieve this is to use selective arsenic diffusion prior to epitaxial growth which results in an epitaxial selective collector type of transistor. Another method of achieving symmetry is the use of a double-sided selective diffusion. In this method, deposits are introduced on opposite faces of the crystal. Masks on both sides are required and alignment is quite critical.

Results so far have been only slightly better than those obtained for the first (0.2-V) effort. Problems arose in handling the thin wafer transistor. Flatness and parallelism have also caused problems in the large area devices. The thin crystal does show promise for smaller devices.

4. Epitaxial-Selective Collector

The epitaxial-selective collector approach involves growing a junction over a buried collector. It is presumed that auto-doping has offset differential penetration, resulting in shallow differential steps. Back voltage of these units were low. It is felt that the problem lies in the basic process technique.

Based on the experience to date, a double epitaxial base approach has been proposed. This process will yield higher absolute gains as well as some symmetry between forward and reverse gains. Geometric symmetry was thought to be the key to low drop. However, because of practical problems in obtaining such symmetry, it is now felt higher absolute gain should help in providing low voltage drop.

Future studies will investigate the use of several matched chips in one package as a simple way to achieve low current density and hence low drop.

C. A Cell for the Direct Observation of Gassing Phenomena at Battery Electrode Surfaces in Low-Gravity Environments, *G. L. Juvinall*

1. Introduction

As part of a continuing JPL program to study the effects of variations of gravitational acceleration on the performance of batteries, the General Electric Research and Development Center is under contract to design and construct a breadboard low-gravity battery test unit. The capabilities of the breadboard unit will include measurement of the limiting current of smooth flooded zinc anodes, measurement of the limiting current and capacity of conventional silver–zinc battery cells, and a photographic study of the formation and behavior of bubbles at the surface of polarized silver and zinc electrodes during flight. The photographic results will be correlated with electrical measurements of capacity and limiting current. Data generated by this program will be directly applicable to the design and construction of spacecraft batteries uniquely suited to future long-term missions contemplated by JPL.

This article discusses the construction and operation of the cell used for the photographic study of battery electrodes.

2. Cell Construction and Operation

The complete test cell assembly and an exploded view are shown in Fig. 6. The cell mounting flange contains the test electrode, which is simultaneously illuminated and observed through the optical window. Continuous electrolyte circulation from the cell mounting flange and the optical window through the second mounting flange and the gas collector is necessary in order to ensure clear optics and gas–liquid phase separation at any time. The cell performance is thus independent of gas pressure or gravity conditions.

Further details of the electrode arrangements in the test cell are illustrated in Fig. 7. The zinc wire reference electrode is positioned close to the test electrode in order to accurately measure its polarization. Also shown is the electrolyte flow path. Electrolyte is constantly swept past the electrode to remove any gas bubbles occurring in the test electrode structure during overcharge and evacuation.

A top view of the cell mounting flange, which contains the test electrode, the counter electrode, and the three leads, is shown in Fig. 8a. An exploded view of the cell mounting flange is shown in Fig. 8b. Inside the cavity of the electrode holder are the silver substrate, with lead, the zinc test electrode (¼-in. disc), a cellophane separator, a disc of nylon mesh, and a disc of expanded nickel. All components are under compression from the expanded nickel grid. The counter electrode holder and counter electrode are also shown.

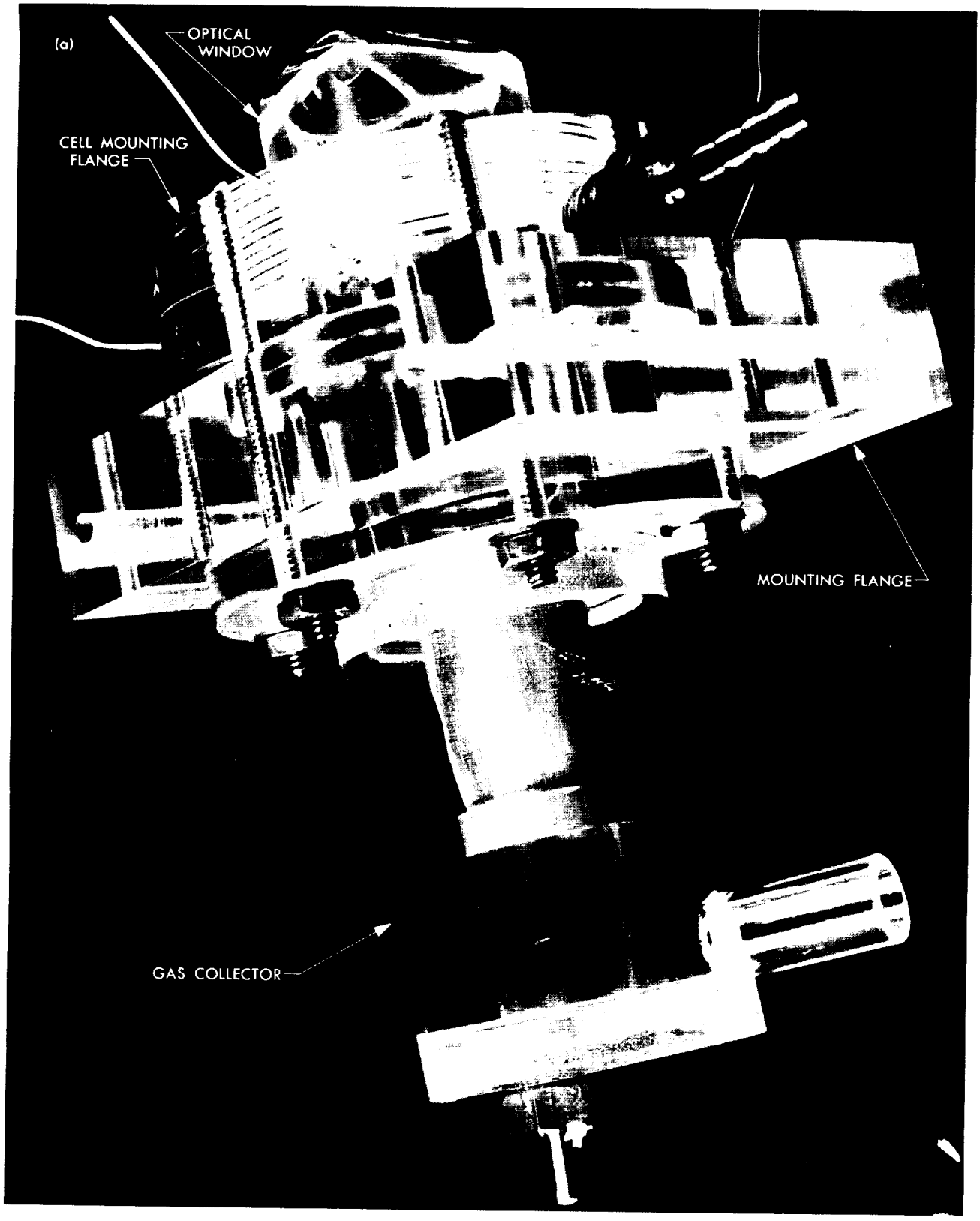


Fig. 6. Test cell: (a) assembled, (b) exploded view

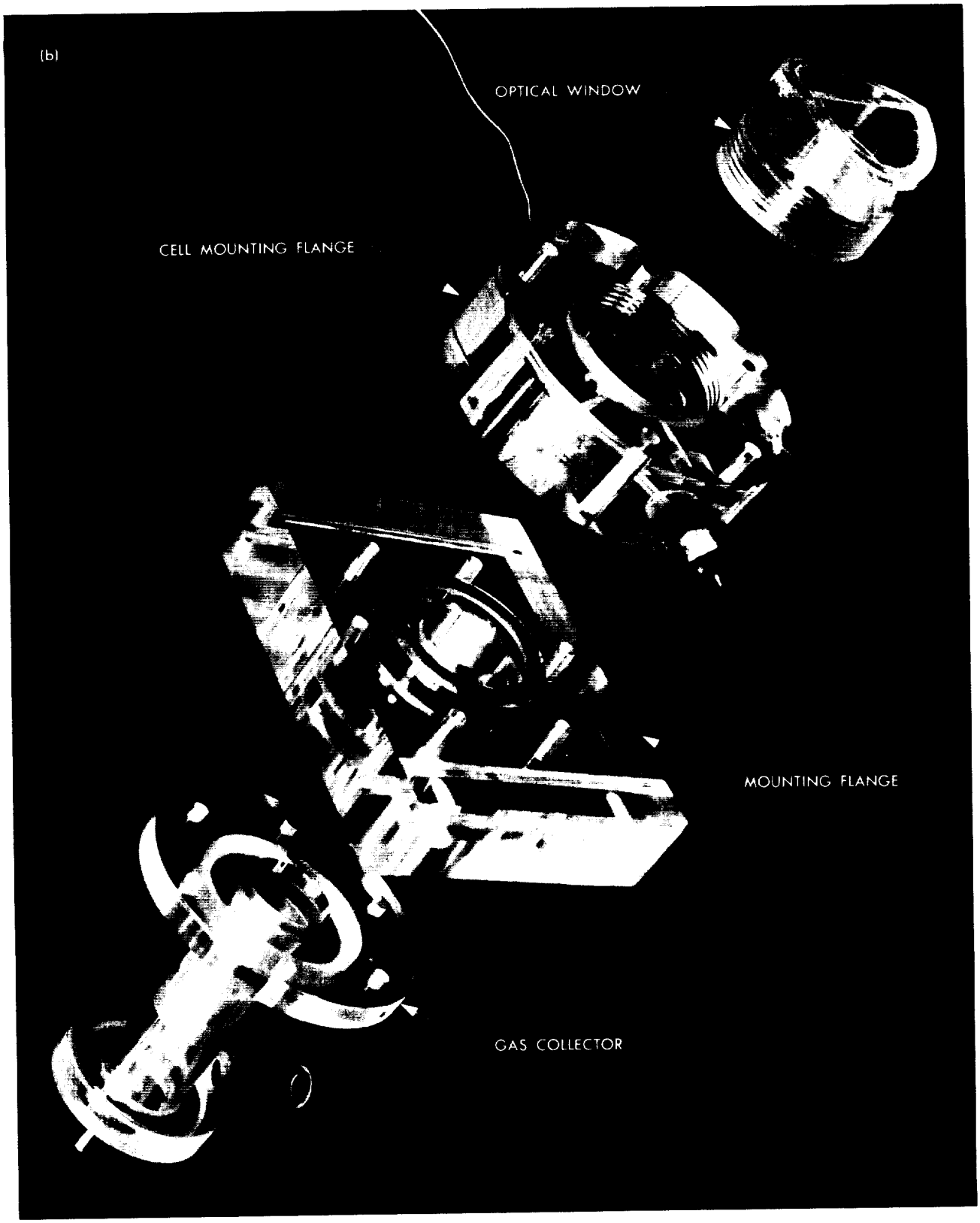

(b)
OPTICAL WINDOW
CELL MOUNTING FLANGE
MOUNTING FLANGE
GAS COLLECTOR

Fig. 6. (contd)

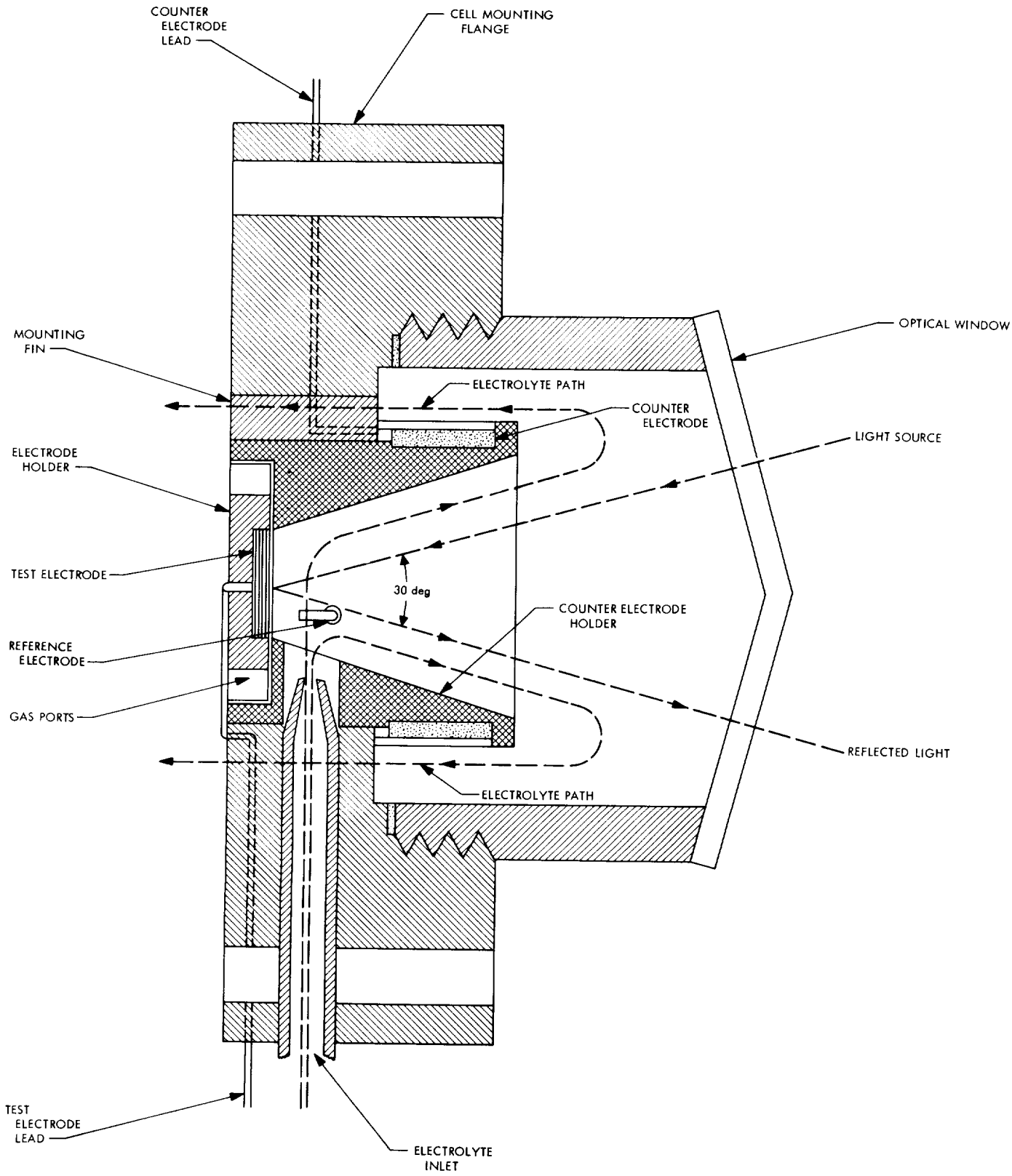


Fig. 7. Arrangement of test electrode and counter electrode in test cell

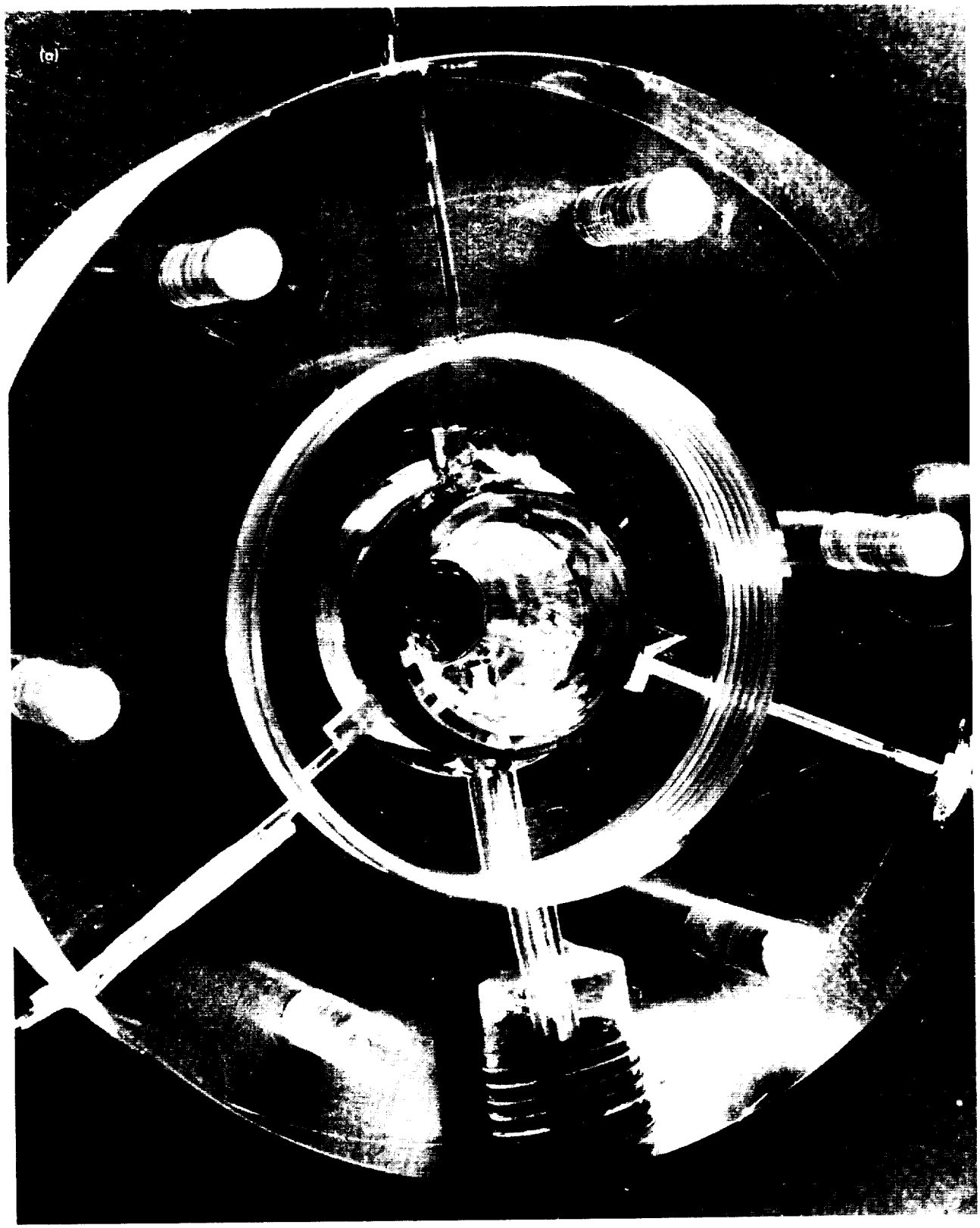


Fig. 8. Cell mounting flange: (a) top view, (b) exploded view

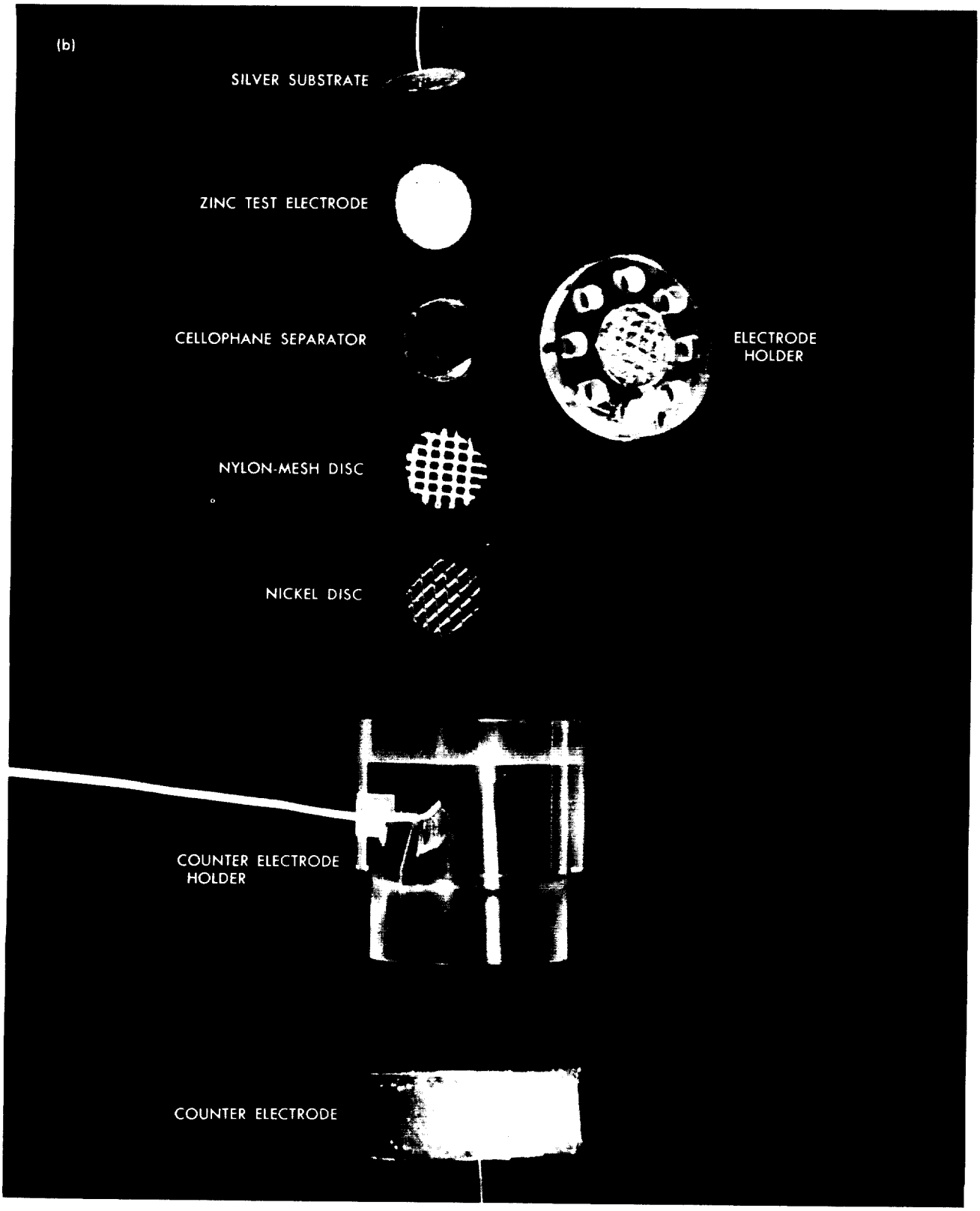

(b)
SILVER SUBSTRATE
ZINC TEST ELECTRODE
CELLOPHANE SEPARATOR
ELECTRODE HOLDER
NYLON-MESH DISC
NICKEL DISC
COUNTER ELECTRODE HOLDER
COUNTER ELECTRODE

Fig. 8. (contd)

During operation, bubbles are induced to form on the surface of the electrode by intentional overcharging. Other bubbles in the system will be removed by the constantly circulating electrolyte (45% KOH). Electrical measurements will be taken during discharge. At the end of discharge, the cell is evacuated and backfilled to ambient pressure. The evacuate-backfill cycle is repeated a second time to reduce the trapped gas volume inside the porous test electrode to an insignificant value.

Aggregates of gas bubbles at the electrode–membrane interface form highly reflective areas, which are easily seen or recorded on films. Such bubble structures have been observed on all test electrodes studied.

3. Conclusions

A successful design has been developed for a cell which will permit the direct observation of bubble formation and behavior at the surface of a battery electrode while the electrical parameters are being measured. The construction of the cell is unique in that the electrode environment closely approaches that in a real battery cell, which contains a compressed pack of electrodes individually wrapped with separator material.

It is anticipated that this cell will become a valuable laboratory tool after fulfilling its intended purpose in the breadboard unit. For example, it affords an opportunity to precisely investigate the efficiency of various gas inhibiting compounds which are added to battery electrodes at the time of fabrication.

D. The Products of the Electrochemical Oxidation of Zinc Battery Electrodes, *G. L. Juvinall*

1. Introduction

Idaho State University is presently under contract to JPL to investigate the reactions pertaining to zinc, cadmium, and silver electrodes. Dr. G. Myron Arcand is the principal investigator on this project. This effort represents part of a continuing study of the fundamental properties and reactions used in alkaline batteries. A major objective of this effort is the acquisition of quantitative data regarding the nature of the electrochemical oxidation products of zinc and cadmium electrodes. The results of another area of study included in this contract, the thermal decomposition of argentic oxide, have been reported previously in SPS 37-49, Vol. III, pp. 122–123. Such information is needed to understand the fundamental electrochemical processes involved in the operation of alkaline spacecraft batteries.

2. Experimental Approach

The method used in the identification of zinc and cadmium reaction products involved the use of tritium as a tracer. If a metal hydroxide is precipitated from a solution containing tritium in the form of either water or hydroxide, the precipitate will show activity. If an oxide is precipitated, it will be inactive. The activity will be proportional to the total hydrogen in the precipitate. No distinction can be made between hydrogen present as hydroxide or water of hydration. The relative amounts of hydroxide and oxide in the precipitate can be determined by measuring the activity of the product. Liquid scintillation counting was accomplished with a Nuclear-Chicago model 703 counting system.

3. Results and Discussion

The relative amounts of $Zn(OH)_2$ calculated from tracer data are shown in Table 3. The high values (above 10%) are probably due to incomplete drying of the precipitate. The results average 2.8 mole-% $Zn(OH)_2$, with σ (standard deviation) 1.3% of the high values discarded. The next lowest value, 9.1%, is 3.2 σ above the average. The statistical probability of a random deviation of this magnitude is less than 0.2%. Thus, tracer analysis indicates that the oxidation product, zinc oxide, contains less than 5% zinc hydroxide. This conclusion is borne out by the results of a thermogravimetric analysis of the reaction product, which is shown in Fig. 9. The initial relative weight appears to be less than 1.000 because the sample was heated to 50°C during temperature equilibration. If the assumption is made that the hydrogen-containing component was removed at temperatures below 120°C, the results compare favorably with the tracer data. In addition, the data in Table 3 show that the nature of the product is unaffected by the form of the electrode.

Table 3. Product of electrochemical oxidation of zinc

Solid $Zn(OH)_2$ formed, %		
Zn sheet anodes	ESB, Inc., porous anodes	Zn-on-Pt anodes
9.1	20	12.0
16.0	2.9	1.1[a]
4.3	5.4	
1.8[a]		
1.7[a]		
2.2[a]		

[a]Reaction in N_2 atmosphere.

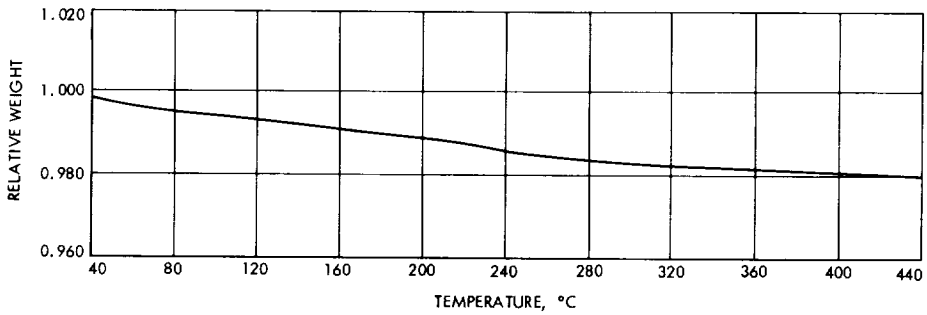


Fig. 9. Thermogram of electrochemical ZnO at $\Delta T = 5$ deg/min

E. Development of the Heat Sterilizable Battery,

R. Lutwack

1. Introduction

The research and development program for a heat sterilizable battery comprises contracts with ESB, Inc., for the Ag–Zn system, with Texas Instruments, Inc., for the Ni–Cd system, and with Eagle–Picher, Inc., for the remotely activated system. The ESB, Inc., contract is supplemented by an in-house effort.

The research and development program for a separator for the heat sterilizable Ag–Zn battery comprises contracts with Monsanto Research Corporation, Westinghouse Electric Corporation, and the Southwest Research Institute. A research and evaluation program is being conducted at JPL.

2. Battery Systems

a. ESB, Inc. (contract 951296). Phase I studies were directed to electrochemical problems and to the design for a high-impact resistant cell. Some of the results are:

(1) A special, very low rate pre-formation charge must be used to prevent excessive gassing at the Ag plate during the formation charge.

(2) A procedure of partial discharges increases and makes more uniform the capacities of cells which are fabricated in the same manner. The partial discharge followed by a recharge increases the cell capacity about 15%.

(3) Capacity data for open-circuit stand at room temperature after 7, 8, and 9 mo have been obtained. Capacity losses per month vary from 0 to 1.3% of the total capacity.

(4) An electrolyte concentration study has shown that, up to 13 cycles, cells with 41% KOH had more uniform capacities than did those using 31% KOH.

(5) A cell pack tightness factor of 2.1×10^{-3} in. per layer produces higher capacity cells than does a larger factor. This has been established up to 11 cycles.

(6) It appears that heat sterilization can be done after the electrical cycling of cells if special precautions are taken to completely discharge the Zn plate before sterilization. Data have been obtained indicating this procedure can be used without creating large gas pressures and without incurring large capacity losses.

These electrochemical and cell design investigations are continuing in an effort to improve the cell performance.

Phase II consists of the design, fabrication, and testing of cells for four types of batteries.

Five AH high-impact resistant cells, in which the plates are reinforced with heavy Ag backbones and the subcover and cover are special designs, were formed into 12- and 14-cell batteries for the capsule systems advanced development program. The sterilized 12-cell primary batteries performed satisfactorily in capsules after impacts of 1200- and 2500-g loads.

In the first tests, the capacity losses due to sterilization of the 80 AH prototype cells for the 2000 W-h battery were low.

The emphasis in the task for the secondary battery is on the Zn plate; teflon-containing plates are being studied.

b. Texas Instruments, Inc. (contract 951972). In the task for electrochemistry studies, which include statistical experiments for characterizing and optimizing the

electrodes, the electrolyte solution, and the separators, it has been shown that:

(1) There are indications from X-ray diffraction, porosity, and pore-size distribution data that sterilization causes morphological and crystallographic changes of the plates, which may affect the solution distribution and the electrochemical efficiency of the cell.

(2) Two sterilization cycles result in about a 30% loss of rated capacity.

(3) The oxygen-free capacity decreases as a result of sterilization although the addition of Co lessens this loss.

(4) The addition of In to the Cd plate has no effect before or after sterilization.

The studies of the effects of sterilization on the properties of the cell components and on the electrical performance of the cell continue.

3. Separators

a. Monsanto Research Corporation (contract 951524). Films prepared from ligand-containing polymers are being investigated. Those which were fabricated on a paper-coating machine from a 31:69 2-vinylpyridine/methyl methacrylate copolymer have good physical properties. Studies of saponification procedures led to the use of a pretreatment with alcoholic KOH solutions to cause partial saponification before the film formation step. Completely saponified films have resistivities of about 10 Ω-cm after sterilization in 40% KOH at 135°C. Evaluation of 100 linear ft of 18-in. wide 1.5-mil-thick film is being done at JPL.

b. Westinghouse Electric Corporation (contract 951525). Composite membranes, which contain a matrix of polypropylene, a binder of polysulfone, and a filler of zirconium oxide, are being investigated. A modified commercial dip-coating procedure and a foil-drying tower have been adapted to fabricate the films. The best specimens, which were prepared with a 3:1 oxide:polysulfone ratio, were shown to have low resistivities and to permit negligible Ag transport. Problems of uniformity and material loss during sterilization remain to be solved. Modifications in the apparatus are being made for the purpose of extending the process to 1-ft-wide tape.

c. Monsanto Research Corporation (contract 951966). Materials made from ethylene/methyl acrylate copolymers, which are crosslinked with a peroxide and hydrolyzed to the salt form in an alcoholic KOH solution, are being investigated. The most promising films were prepared from a copolymer composition of 45:55 for acrylate:ethylene. After consecutive hydrolyses for 24 h at 65°C in alcoholic KOH and then in aqueous KOH solution, these films survived sterilization and had resistivities less than 20 Ω-in. Samples have been submitted to JPL for evaluation.

d. Southwest Research Institute (contract 951718). This is a study of the parameters involved in the preparation of separator materials from polyethylene, which is modified by crosslinking with divinylbenzene and grafting with acrylic acid using Co^{60} radiation. The parameters studied and the results obtained are:

(1) Acrylic, methacrylic, oleic, and maleic acids were used for grafting. The best films were prepared using acrylic and methacrylic acids.

(2) By varying the washing procedure, it was shown that the dimensional changes which occur during sterilization are dependent upon the temperature of the wash. The reasons for this dependency are not known.

(3) Films with higher K contents are produced when the grafting solution contains lower chain terminator concentrations.

(4) The use of Ni, Co, and Zn naphthenates reduces homopolymerization and inhibits grafting.

(5) Films are grafted more uniformly under a N_2 atmosphere although the presence of O_2 has no apparent effect on the crosslinking.

A scale-up of the basic procedure is also being developed. Southwest Research Institute supplies JPL with the separator material used in the Ag–Zn cell development program.

F. RTG Test Laboratory, *R. G. Ivanoff*

1. Introduction

Future mission studies conducted by JPL indicate the need for a solar-independent power source for mission opportunities beginning in 1972. The most probable candidate for this category of power source is the radioisotope thermoelectric generator (RTG). A key problem regarding RTGs is one of incorporating the device into spacecraft designs. Major technologies which must be evaluated are those related to the compatibility of the isotope power source and its inherent radiation field with space science experiments and other sensitive subsystems.

Also, assurance must be provided through life and performance testing to obtain information on the behavior of the device as a power generator.

To support the above requirements, JPL has begun the design and development of an RTG test laboratory. The test laboratory will play a major role in developing technology required to insure integration of RTGs on advanced planetary vehicles. The test laboratory will amplify and complement the capability now present at JPL by providing facilities and equipment for the evaluation and testing of fueled SNAP[1] generators and other isotopic sources.

2. Operations

The test facility is designed to run life and performance tests on both fueled and electrically heated RTGs. These tests include: parametric tests at preselected values of power output, voltage output, and hot junction temperature; short-circuit current tests; power system tests; thermal-vacuum tests; life tests; simulated mission profile tests; and evaluation tests on electronic equipment to assess radiation damage at distances and relative locations simulating those in a spacecraft.

The test laboratory is designed to conduct a program to establish interface requirements. Test objectives of such a program include: (1) identification of the mechanism of science-instrument and neutron- and photon-radiation interference, the level of those interferences, and the nuclear spectral components primarily responsible for the interference effects; (2) prediction of the statistical or transient response of science instruments proposed for typical particle and flux measurement space experiments; and (3) determination of the statistical transient response of as many science instruments as are physically obtainable, in the environment of expected neutron and photon mixed beam dose rates. To identify further the mechanism of the nuclear radiation and science instrumentation interferences, the measurements will be correlated with the responses predicted from theoretical studies.

Other more general objectives of the test laboratory are:

(1) Provide the spacecraft designer and experimental scientists with accurate data on the interactions between the power supply and the spacecraft and/or scientific experiments.

(2) Create an installation capable of supplying and dispersing knowledge of the interrelation/integration between the isotope power sources and different scientific experiments/instruments, components, and spacecraft configurations.

(3) Obtain performance data that is unavailable when using electrical heating.

(4) Provide NASA and other government agencies or facilities an independent and accurate report on the performance of isotope-fueled power sources.

(5) Familiarize JPL personnel with the handling of radioisotope-fueled devices.

Fueled SNAP generators for the laboratory tests will be obtained on a loan basis from the Atomic Energy Commission (AEC) and will remain under AEC compliance.

3. Task Status

Equipment design and construction details for the RTG test laboratory have been established. Preliminary test flow plans are completed. Security and radiological safety requirements have been established with AEC and NASA. Certain special laboratory equipment has been procured.

G. Analog Voltage to Duty Cycle Generator,[2]

A. I. Schloss

In all switching types of regulators, there is a circuit that converts a dc error signal to a time-ratio switching signal. The ratio of switch-on time to cycle time is called a duty cycle. The greater the duty cycle the greater will be the dc output voltage of the switching regulator. The circuitry that accomplishes this function operates at a low power level and lends itself to fabrication as an "integrated" or "micro" circuit. Integrated circuits reduce the number of required interconnections relative to those required in a discrete component approach for the same function. Reliability and long life are, therefore, greatly enhanced. The purpose of this effort is to produce a "universal" duty-cycle generator in integrated form. It is universal in that it is designed in a way to be easily adaptable to many different switching regulator designs. The regulator may vary as to power and voltage level and may be of either the single- or double-ended design.

[1]Systems for nuclear auxiliary power.

[2]Contractor for this project is Westinghouse Research Laboratories, Pittsburgh, Pa.

Besides the increase in reliability offered because of the integrated approach, there is also the possibility of increasing reliability using high-level redundancy. The size and weight of the integrated circuit permits relatively complex (high-level) circuits to be used redundantly in a practical way.

A discrete component equivalent of the duty-cycle generator has been completed to demonstrate circuit performance. The circuit performed to specification. Masks were made to fabricate the monolithic circuit. It has been determined that the best compromise between yield and the number of external connections would occur when four chips are used. The four chips are to be bonded to a single heat sink. Layout of the four chips is such that if yield is good it will be possible to produce the generator on one piece of silicon.

It has been demonstrated that all functional parts of all chips have at one time or another worked; however, to date there have not been four completely functional chips with which to make one generator. Low yield is aggravated by the limited production of this research effort. Much of the effort so far has been in perfecting fabrication techniques. A complete generator should be available for test within the next three reporting periods.

VI. Spacecraft Control

GUIDANCE AND CONTROL

A. Development of an Approach-Guidance Optical Planet Tracker, *F. R. Chamberlain*

1. Introduction

Competing analytical theories for optical approach-guidance orbit determination are under development by a number of organizations, including JPL. While some studies emphasize long-duration measurements several weeks prior to encounter, others place heavy emphasis on measurement of planet angular diameter during the last 24 h. All theories characterize accuracy limitations in terms of angular measurement bias and drift. Figure 1 illustrates a typical case where 1-σ tolerances of 4 arc-sec drift and 40 arc-sec bias are assumed.

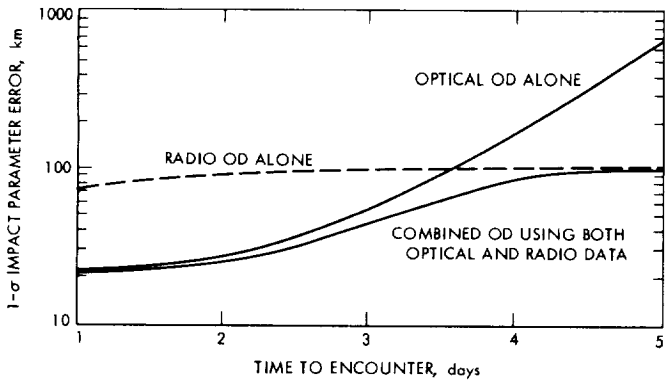


Fig. 1. 1-σ impact parameter errors vs time

Development of an optical planet tracker was initiated by JPL in April 1966; detailed design and fabrication were subcontracted to Electro-Optical Systems, Inc. Program plans during FY 67 included a flight-feasibility demonstration on *Mariner* Mars 1969 as an engineering experiment. The initial efforts were reported in SPS 37-42, Vol. IV, pp. 48–49. Preliminary and detail design were accomplished, followed by fabrication of flight-configuration planet tracker prototypes and precision equipment for evaluation of hardware performance. Parts for flight hardware were bought and screened, and a detailed design review and exploratory environmental testing were completed.

A revised program for ground hardware evaluation and engineering development has now been implemented at JPL. Facilities have been modified to provide a temperature-stabilized and seismically isolated environment for limited demonstrations of feasibility. Studies continue in preparation for either another flight-experiment opportunity or a mission-critical application where optical approach guidance is mandatory for reasons of required orbit determination accuracy.

2. Data Correlation Concepts

The measurements made by an optical planet tracker will exhibit the effects of attitude motion and, if the planet tracker's measurements are corrected by means of attitude-control system data, all attitude-control system errors in

bias and drift will carry over into orbit determination calculations. In addition to the sum bias and drift of individual attitude-control sensors, each alignment tolerance is a bias; any alignment drift during a measurement period is equivalent to an instrument drift.

As a result of attitude-motion considerations, auxiliary attitude-motion reference sensors with arc second accuracies are required; incorporation of these sensors' optical trains into the planet-tracker chassis is highly desirable for alignment reasons. The possibilities of superimposing planet and reference-body images (so that correlation reduces to precision opto-mechanical or electro-optical scanning) are very attractive and are under evaluation.

Past development of auxiliary attitude-motion sensors for approach guidance is limited to a miniaturized narrow-angle sun sensor of conventional shadow-bar and cell configuration. Ultra-stable excitation circuitry, high-resolution thermal sensors, and exhaustive calibration will be required to achieve a repeatable analog output. Analog-to-digital conversion will require either additional approach-guidance subsystem electronics complexity or stringent requirements for the analog-to-digital circuits in a telemetry subsystem.

3. Planet Tracker Concept

The planet tracker utilizes an electrostatic image dissector as its sensor, and scans the image of a 10 × 10-deg field of view over the sensor's faceplate by means of orthogonal pairs of counter-rotating optical wedges. It searches for, locks onto, and tracks the geometrical center of gibbous-phase planets ranging in size from 1–0.05 deg in angular diameter.

Figure 2 illustrates the layout of the tracker opto-mechanical train and shows locations of the electronic modules. After the field of view is corrected by a foreprism and scanned by optical wedges, it is focused through a 24-in. folded optical path onto the image dissector faceplate. The electronics mechanization provides the logic and signal processing for search, acquisition, and planet tracking; it scans and derives signals from the image dissector tube, controls wedge positions through servo-mechanism drives, and reports the angular positions of wedge encoders to telemetry circuits through a data converter.

In the search mode of operation, a raster pattern is scanned across the 10 × 10-deg field of view. The image area interrogated by image dissector scanning is a circle 1 deg in diameter, and a raster interval slightly larger than

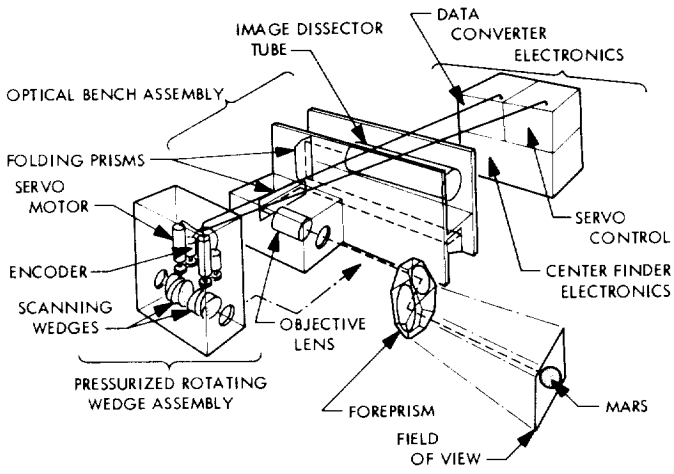


Fig. 2. Optical train diagram

0.2 deg is used to provide extensive overlapping. The raster is scanned until image dissector signals indicate the presence of a planet. "Planet-detect" circuitry then cues in an acquisition mode to approximately center the planet.

In the centering mode of operation, orthogonal bipolar error signals are derived from a rotating spike (starburst-shaped) scanning pattern and are used to approximately center the planet. Once this preliminary tracking null is obtained, transition to a geometrical tracking mode is carried out.

In the geometrical tracking mode, an image dissector scan pattern centered about the electron aperture is used that scans only the illuminated limb of the planet and develops error signals that not only recenter the planet, but also adapt the scan radius for changes in planet angular diameter. In signal processing, a memory is developed of peak planet brightness for each scan orientation and thresholding for the generation of pulse-width information is based on a fixed proportion of the memory value. This method of signal processing largely eliminates albedo-related amplitude errors.

A data converter periodically interrogates the two 10-bit encoders for wedge (field of view) positions and converts the 20 bits of parallel format information, plus 8 engineering bits, into 4 telemetry words of 7 bits each in serial format.

4. Design Criteria and Characteristics

The existing approach-guidance planet tracker weighs 15 lb, consumes less than 8 W of power, and occupies an envelope with a maximum dimension of 12 in. It is

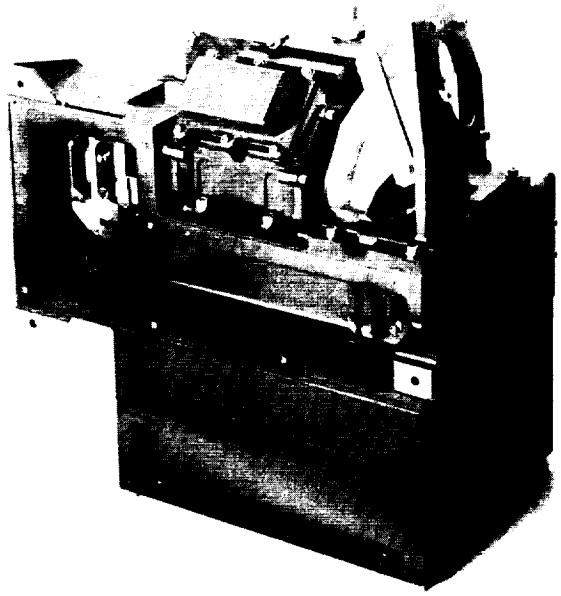


Fig. 3. Approach-guidance planet tracker

designed to survive temperatures ranging from 0–120°F. Figure 3 is a photograph of a planet tracker with the optical train cover removed.

Repeatability of the optical gimbals is specified at 4 arc-sec; an alignment stability of 2 arc-sec over 4 days is a design goal. All aspects of the opto-mechanical design involve detailed attention to the effects of temperature variation, thermal gradients, and mechanical errors; gear train and stress analyses were performed in the selection of gear configurations, anti-backlash methods, and bearings.

A closed compartment is required for the mechanisms, primarily to avoid particulate contamination. While pressurization is used to avoid the complexity of venting devices, gear materials are selected for resistance to cold welding phenomena and vacuum-qualified lubricants are employed. All materials and components to be enclosed by the pressurized compartment have been subjected to a thorough contamination test to avoid the possibility of some high vapor pressure substance (e.g., residue from a cleaning process in a motor or encoder) redepositing on cold optical surfaces and degrading performance. Pressure switches are installed in the pressurized compartment for optical calibration correction in case of pressure loss. (Loss of all pressure results in a 10 arc-sec increase to wedge deviation at the edges of the field of view.) Figure 4 is a view of the scan subassembly with one of its pressure covers removed.

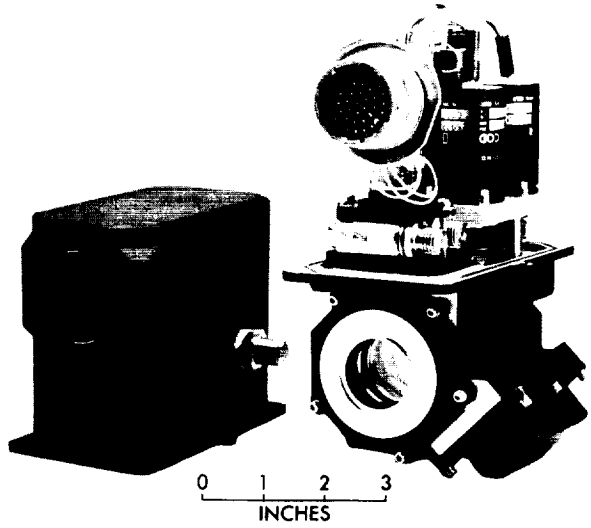


Fig. 4. Scan subassembly

The image-forming subassembly of the approach-guidance planet tracker contains a 24-in. focal-length triplet telescope lens, 3 porro prisms to fold the optical path, and the image dissector tube. Ahead of the objective lens are the 4 achromatic wedges (two elements per wedge), the two pressurized compartment windows, and a foreprism. Scattering, absorption, and reflection losses in the thirteen elements sum to a transmission coefficient of 0.670 that, when used with a 20% margin for degradation and an objective lens diameter of 1.10 in., yields an effective clear aperture of 0.531 in.^{2}

Design of the planet-tracker objective lens results in a nearly diffraction-limited image with an Airy disc diameter angular subtense of 9 arc-sec. The achromatic wedges each contribute a difference of deviation such that when weighted for spectral power, the new wedge dispersions at the edge of the field of view are about 8 arc-sec. The limiting element in system resolution is the image dissector tube, which has an electron aperture diameter of 0.004 in. and an electron spot dispersion of the same dimension. The 69 arc-sec subtended by the electron aperture and electron dispersion serve to dominate all modulation transfer function descriptions.

In the electronics mechanization, wide use is made of microcircuits for logical operations and signal processing; field-effect transistors are used for switching functions in the generation of digital scanning signals and analog memories. Electronics packaging techniques emphasize

the use of welded cordwood modules for discrete components and stick modules for integrated circuits; circuit boards were used in the immediate vicinity of the image dissector tube for high-voltage power supplies and the video preamplifier.

5. Conclusion

Problems for continued study include the correlation of tracking data for the planet with that for celestial reference bodies (within an extremely tight error budget). The difficulties of sensor integration into a spacecraft system, and software integration with mission operations, may not be subject to satisfactory short-term solutions; a flight experiment prior to a mission-critical application is the preferred approach.

B. Strapdown Inertial System Analysis, *G. Paine*

1. Introduction

The Strapdown Electrically-Suspended Gyro Aerospace Navigation (SEAN) system is a developmental inertial-navigation system described in SPS 37-36, Vol. IV, pp. 55–59. The system employs two-position electrostatic gyroscopes (ESGs), and three digital velocity meters (DVMs) as inertial sensors. Each ESG has three pickoffs to provide two, or all three, direction cosines between the case and the rotor spin axis. The ESGs have been described previously (SPS 37-36, Vol. IV, pp. 51–55). The ESG drift model and its development is described in *Section D*. This article discusses the systems-analysis effort and the resultant software. Due to their number, detailed equations are not included here.

The initial analysis resulted in a set of equations used in the theoretical part of the error analysis. Subsequently, this set was expanded and embodied in several system simulators. Two of these programs, described here, are written in Fortran IV and can be executed on an IBM 7094–7044 computer installation. The software developed for the digital flight computer that processes the outputs of the inertial sensors is flow-charted. The flight equations were a direct outgrowth of the system-simulator equations.

The interconnection between the theoretical error analysis and the two simulators described can be seen in Table 1. A variety of error sources are listed together with marks identifying which method of analysis was used to determine the magnitude of their effect.

The concept of testing system behavior by simulation prior to mechanization of the flight computer software has been justified and has resulted in a higher degree of confidence in the final software because of the thoroughness of its checkout. Without the flexibility and accuracy of a large-scale computer, a complete checkout would be considerably more difficult.

Table 1. System error sources and analysis methods

Error source	Confirmation source[a]		
	T	1	3
Attitude errors			
Initial attitude (alignment)			
Uncompensated gyro drift	D		D
Normal gyro data noise	D		D
Bad gyro counts	I		I
Gyro count resolution	D		D
Accelerometer data resolution	D		D
Accelerometer errors	I		I
Attitude measurements			
Normal gyro noise	D		D
Bad gyro counts	I		I
Gyro pattern misalignment	I		I
Gyro misalignment	D		I
Vehicle turning rate	D		D
Averaging gyro direction cosines	D		D
Gyro drift (two gyros)	D		D
Gyro drift sine term	D		
Velocity errors			
Hardware			
Accelerometer bias	D	D	
Accelerometer scale factor	D	D	
Accelerometer misalignment	D	D	
Software			
Computed gravity	D		
Systematic with rotation rate	D	D	
Accelerometer resolution with rotation rate	D	D	
Systematic with oscillation	D	D	
Accelerometer resolution with oscillation	I	D	
Systematic with vibration			D
Position errors			
Hardware			
Altimeter bias	D	D	
Altimeter resolution and noise	D	D	
Software			
Initial position	D	D	
Computation and integration		D	D

[a]T = theoretical
1 = simulator 1
3 = simulator 3
D = Detailed investigation
I = Incomplete investigation

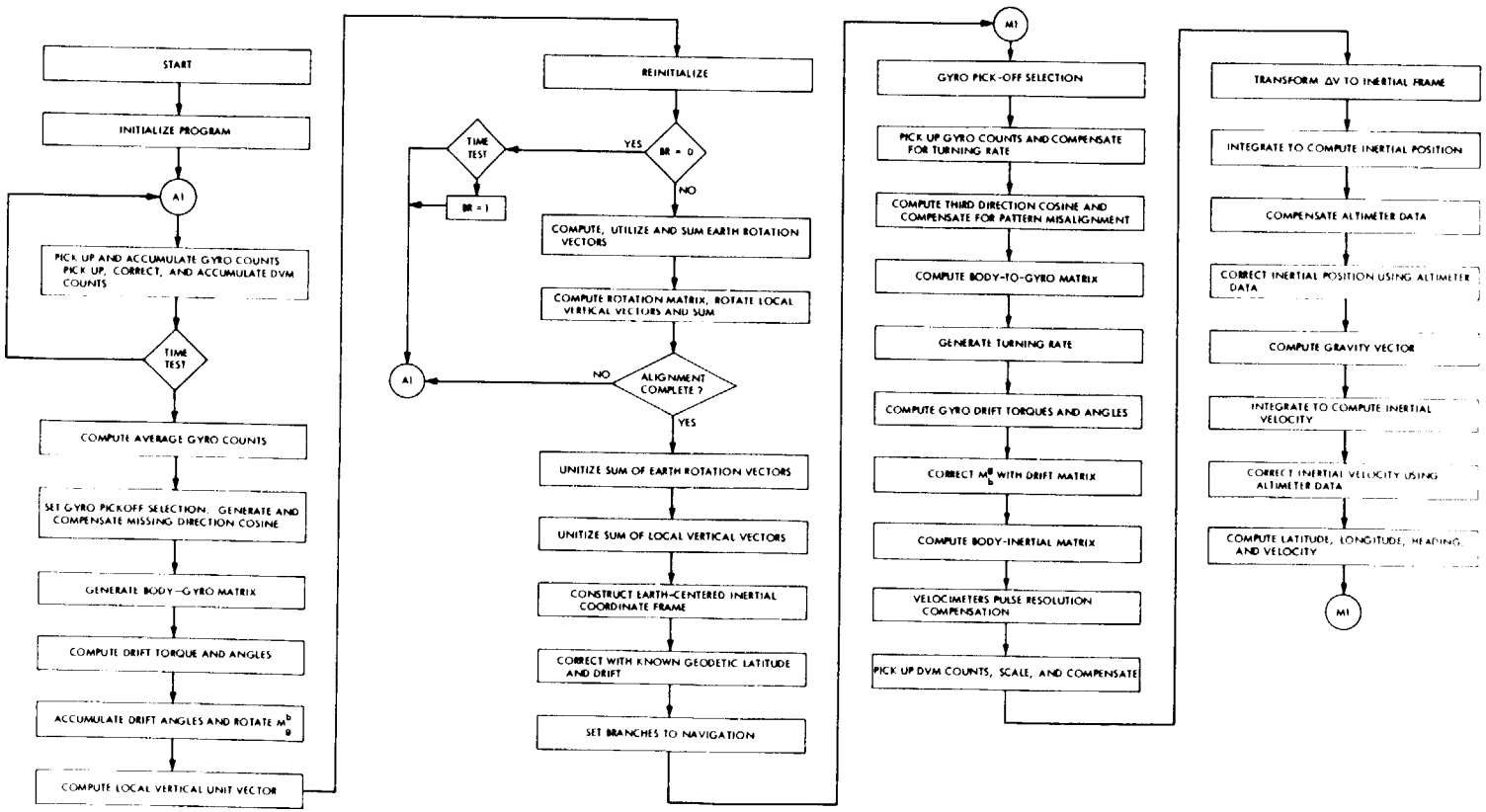

START
INITIALIZE PROGRAM
A1
PICK UP AND ACCUMULATE GYRO COUNTS PICK UP, CORRECT, AND ACCUMULATE DVM COUNTS
TIME TEST
COMPUTE AVERAGE GYRO COUNTS
SET GYRO PICKOFF SELECTION. GENERATE AND COMPENSATE MISSING DIRECTION COSINE
GENERATE BODY–GYRO MATRIX
COMPUTE DRIFT TORQUE AND ANGLES
ACCUMULATE DRIFT ANGLES AND ROTATE M_g^b
COMPUTE LOCAL VERTICAL UNIT VECTOR
REINITIALIZE
BR = 0
YES
NO
TIME TEST
BR = 1
COMPUTE, UTILIZE AND SUM EARTH ROTATION VECTORS
COMPUTE ROTATION MATRIX, ROTATE LOCAL VERTICAL VECTORS AND SUM
A1
NO
ALIGNMENT COMPLETE ?
YES
UNITIZE SUM OF EARTH ROTATION VECTORS
UNITIZE SUM OF LOCAL VERTICAL VECTORS
CONSTRUCT EARTH-CENTERED INERTIAL COORDINATE FRAME
CORRECT WITH KNOWN GEODETIC LATITUDE AND DRIFT
SET BRANCHES TO NAVIGATION
M1
GYRO PICK-OFF SELECTION
PICK UP GYRO COUNTS AND COMPENSATE FOR TURNING RATE
COMPUTE THIRD DIRECTION COSINE AND COMPENSATE FOR PATTERN MISALIGNMENT
COMPUTE BODY-TO-GYRO MATRIX
GENERATE TURNING RATE
COMPUTE GYRO DRIFT TORQUES AND ANGLES
CORRECT M_b^g WITH DRIFT MATRIX
COMPUTE BODY-INERTIAL MATRIX
VELOCIMETERS PULSE RESOLUTION COMPENSATION
PICK UP DVM COUNTS, SCALE, AND COMPENSATE
TRANSFORM ΔV TO INERTIAL FRAME
INTEGRATE TO COMPUTE INERTIAL POSITION
COMPENSATE ALTIMETER DATA
CORRECT INERTIAL POSITION USING ALTIMETER DATA
COMPUTE GRAVITY VECTOR
INTEGRATE TO COMPUTE INERTIAL VELOCITY
CORRECT INERTIAL VELOCITY USING ALTIMETER DATA
COMPUTE LATITUDE, LONGITUDE, HEADING AND VELOCITY
M1

Fig. 5. SEAN flight computer program flow chart

2. System Simulators

In order to verify the theoretical system analysis and to investigate areas that are not amenable to theoretical studies, several system simulators were written. Each of these allows computer simulation of the SEAN system performance. Each simulator is divided into (1) a driver that simulates the inertial sensors over some test flight path, and (2) a navigation portion that simulates the behavior of the flight computer software. The generation of accurate inertial sensor data consumes the largest portion of computer time during simulation. To simulate the system efficiently, some of the capabilities of the driver are provided as options so that the associated computations will be skipped if the options are not being used.

a. Navigation simulator 1. This simulator simulates the SEAN system minus the gyros. The driver provides DVM outputs and an attitude matrix to the navigation simulator. These are provided for a constant latitude, constant velocity, and constant altitude flight path. In addition, the driver can simulate a combination of roll, pitch, and yaw angles coupled with constant and sinusoidal roll, pitch, and yaw rates. The navigation simulator processes the DVM outputs with the attitude matrix provided to give delta velocity quantities in an earth-centered inertial frame. The delta quantities are then integrated to obtain position, velocity, latitude, and longitude.

This simulator was used to verify (1) the predicted effects of accelerometer bias, scale factor, misalignment, and resolution with constant and sinusoidal turning rates, (2) the effects of systematic software errors caused by vehicle turning rates, (3) the accuracy of the integration procedure used, and (4) the effects of altimeter bias, resolution, and noise.

b. Navigation simulator 3. This simulator simulates the complete SEAN system. The driver provides DVM and ESG outputs to the navigation simulator for a constant latitude flight path. The driver can simulate the same combinations of constant and sinusoidal turning rates as simulator 1. This driver can also simulate the reversal of the flight path, a change of altitude, and a flight path where the aircraft altitude is perturbed by sinusoidal variations. The navigation simulator processes the DVM and ESG outputs to generate position, velocity, latitude, and longitude. The same earth-centered inertial frame is used here as in the navigation portion of simulator 1. In addition, the initial alignment program is simulated.

This simulator was used to check out the final equations used in the flight computer program (subsequently flow-charted) and to verify the predicted effects in alignment of gyro drift, noise, and resolution, and of accelerometer resolution. In navigation, the effects of gyro drift, noise, and resolution, and the effects of averaging gyro counts, were verified. Wherever possible, simulator 3 was checked against simulator 1.

3. Flight Computer Program

The equations generated and verified by the system simulators have been assembled into a final form for definition of the flight computer program described below.

The flow chart of the flight program (Fig. 5), shows how the inertial sensor outputs are processed to generate the initial alignment data, and then, during navigation, to generate velocity, heading, latitude, and longitude information. The alignment process is described in *Section C*.

The outputs from the inertial sensors are sampled and processed at discrete intervals. Between samplings, the outputs from each sensor are accumulated. An external clock maintains the synchronism between the sampling of the inertial measurement unit outputs and the execution of the computer program.

For the sake of brevity, several loops that determine the frequency of some of the corrections in navigation have been omitted, as have the variables that determine the intervals of summation and averaging in alignment.[1]

C. Strapdown Inertial System Alignment Technique, *B. R. Markiewicz*

1. Introduction

A Strapdown Electrically-Suspended Gyro Aerospace Navigation (SEAN) system, presently under development at JPL, uses 3 accelerometers and 2 two-degree-of-freedom attitude gyros mounted rigidly to a common frame defined as the inertial measuring unit (IMU). The IMU is, in turn, rigidly connected (through vibration isolators) to a vehicle body. The IMU supplies the necessary physical measurements to the airborne computer to enable a continuous computation by the navigation program of vehicle position in terms of latitude and longitude.

In addition to performing the navigation function, this system is self-aligning since no external inputs are required. The only restriction on the IMU orientation is

[1]Markiewicz, B. R., *SEAN Navigation Equations for the Alert Computer*, April 30, 1968 (JPL internal document).

that it remain constant with respect to the rotating earth during the alignment period. Although this article does not include any details of the navigation program, since that function is not essential to an understanding of the alignment process, a description of the coordinate systems used in the navigation system will facilitate the present discussion.

2. System Coordinates

The SEAN system uses the three ortho-normal coordinate frames depicted in Fig. 6. The main computational frame is the earth-centered equatorial inertial (ECI) frame with coordinates x, y, and z. All accelerometer data and computed gravitational acceleration are transformed to this inertial frame, which is established at the beginning of navigation (t_0). The gyro frame (g_1, g_2, and g_3) is inertial except for gyro drift and is established from the two gyro spin vectors (which are not orthogonal) by selecting one spin vector as the reference coordinate and orthogonalizing. Finally, the IMU body axes are defined as colinear with the accelerometer triad except for small known misalignments.

It is important to note that the gyro attitude information obtained during every sampling interval (computation cycle time) supplies the orientation of the body system, with respect to the gyro system, in terms of a transformation matrix (M_g^b). This changing matrix is the essence of the alignment process.

3. Alignment Procedure

The basic reason for the alignment process is to establish the orientation of the gyro spin vectors with respect to the computational ECI coordinates. This is accomplished by determining the matrix, M_I^g, which defines the transformation from the gyro frame to the ECI frame. Once this transformation is determined, it remains constant except for gyro drift (an added compensation). The

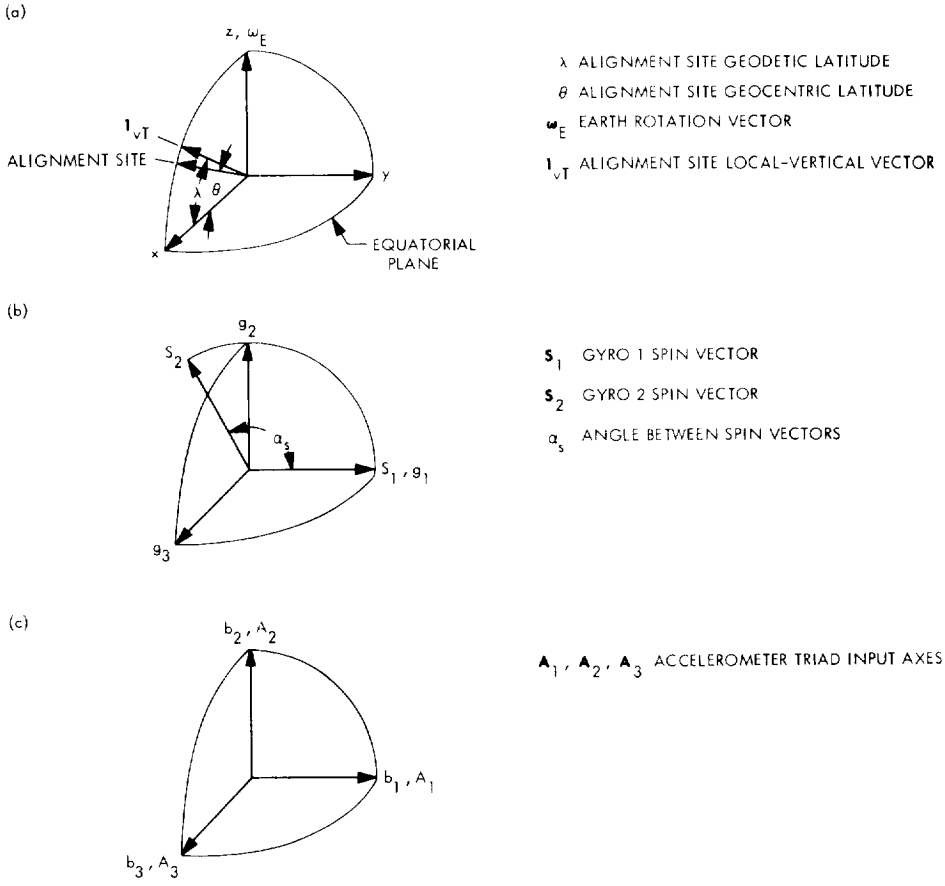


Fig. 6. SEAN system coordinate frames: (a) earth-centered equatorial coordinates, (b) gyro coordinates, (c) body coordinates

subsequent transformation from body coordinates to ECI coordinates is easily obtained as

$$M_I^b = M_I^g M_g^b$$

using the continuously-computed M_g^b.

Deriving M_I^g requires (1) the unitized earth-rotation vector ($\mathbf{1}_{\omega Eg}$), and (2) the alignment-site geodetic vertical vector ($\mathbf{1}_{rTg}$). Both vectors are in gyro coordinates as indicated by the subscript g. The earth-rotation vector is obtained from gyro-attitude information using successive values of M_g^b obtained over an interval (Δt_2).

$$\mathbf{A} = M_g^b (M_g^b)^T_{i-\Delta t_2} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\boldsymbol{\omega}_{Eg} = \begin{bmatrix} \omega_{Eg1} \\ \omega_{Eg2} \\ \omega_{Eg3} \end{bmatrix} = \frac{1}{2\Delta t_2} \begin{bmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{bmatrix}$$

The local-vertical vector is derived from the accelerometer thrust velocity data ($\mathbf{v}_{Tb}$)

$$\mathbf{1}_{rTb} = \frac{\mathbf{v}_{Tb}}{|\mathbf{v}_{Tb}|}$$

and is then transformed to gyro coordinates and rotated about the earth-rotation vector to an inertial orientation coincident with the alignment site at the end of the alignment period.

Many of these vectors are computed and summed over the alignment period. The final unitized mean values are used to compute M_I^g.

$$\mathbf{Z}_g = \mathbf{1}_{\omega Eg}$$

$$\mathbf{Y}_g = \mathbf{1}_{\omega Eg} \times \frac{\mathbf{1}_{rTg}}{|\mathbf{1}_{\omega Eg} \times \mathbf{1}_{rTg}|}$$

$$\mathbf{X}_g = \mathbf{Y}_g \times \mathbf{Z}_g$$

$$M_I^g = \begin{bmatrix} g_{1x} & g_{2x} & g_{3x} \\ g_{1y} & g_{2y} & g_{3y} \\ g_{1z} & g_{2z} & g_{3z} \end{bmatrix} \equiv \begin{bmatrix} x_{g1} & x_{g2} & x_{g3} \\ y_{g1} & y_{g2} & y_{g3} \\ z_{g1} & z_{g2} & z_{g3} \end{bmatrix}$$

This matrix is established at the end of alignment, at which time navigation begins. Thus, the inertial x axis (Fig. 6) is on the alignment-site meridian at zero navigation time, provided the computed M_I^g has no errors.

4. Error Sources

Errors in the alignment-computed M_I^g are the result of errors in the computed $\mathbf{1}_{\omega Eg}$ and $\mathbf{1}_{rTg}$. Attitude deviations in these vectors occur from the sources given below.

a. Earth-rotation vector error sources.

(1) Noise on the gyro attitude data.

(2) Gyro data resolution.

(3) Uncompensated misalignments.

(4) Uncompensated gyro drift.

b. Local-vertical vector error sources.

(1) Uncompensated accelerometer bias.

(2) Accelerometer scale factor error.

(3) Accelerometer data resolution.

(4) Uncompensated accelerometer misalignments.

(5) Uncompensated gyro drift.

Although all of these error sources have been analyzed to a varying degree (and their effects are known), the noise on the gyro data is of primary concern and has shaped the alignment program to its present form.

5. Simulation Results

The accuracy of the alignment program was checked by supplying it with simulated gyro and accelerometer data obtained from a driver simulation developed for the navigation program. This driver program, combined with the navigation program and the alignment program, enabled a complete simulation of the alignment process for any alignment-site latitude and IMU orientation.

The effect of noise on the gyro-attitude data was obtained using a random number generator. To obtain an error characteristic independent of other error effects, the value of noise used was much larger than the realistic value. The results are plotted in Fig. 7 for a 30-min alignment period. The error characteristics, shown as a function of IMU azimuth orientation, are the statistical mean values of the misalignments in M_I^g. Azimuth error and IMU orientation imply a rotation about the local geocentric vertical, while longitude and latitude errors are rotations of M_I^g which result in longitude and latitude

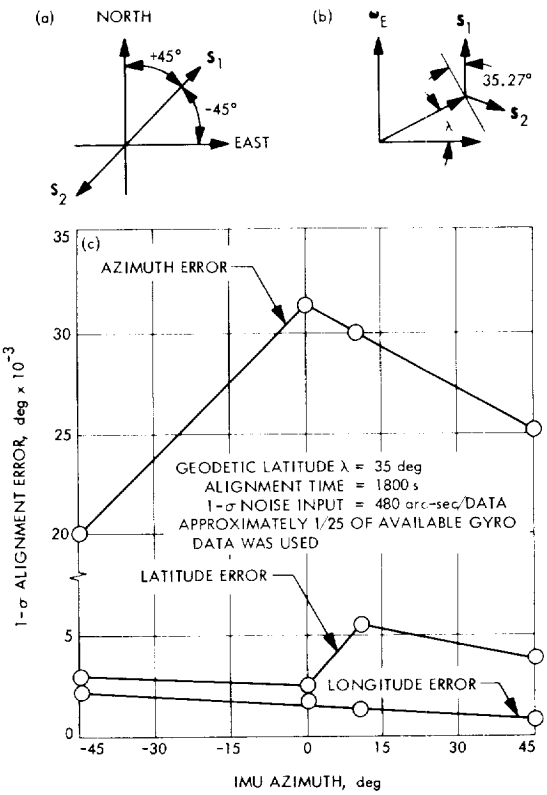


Fig. 7. Alignment errors due to gyro data noise: (a) IMU azimuth in local-horizontal plane, (b) gyro spin vector orientation, (c) 1-σ alignment error vs IMU azimuth

orientation errors, respectively, in the computational ECI coordinate frame.

The fact that only 4 azimuth points were taken for the data plotted in Fig. 7, and only 10 simulation runs were made for each point, accounts for the irregular error characteristic. Using this data, the final alignment error due to noise on the gyro-attitude data can be approximated by considering that the noise magnitude simulated is about 6 times the expected noise, and that only 0.04 of the available gyro data was used to obtain this data. Using all of the gyro data, and a realistic 1-σ noise magnitude of 80 arc-sec on each gyro data set (instead of the 480 arc-sec actually used), results in an error reduction factor of 30 that applies directly to the azimuth error and, to a much lesser degree, to the latitude error. Using an azimuth error of 0.03 deg (Fig. 6), a realistic 1-σ error becomes 0.001 deg, or 3.6 arc-sec.

It has been previously mentioned that no external inputs are required for the alignment process, but alignment-site latitude (although not essential) is assumed known and was used in the program to perform a correction. This accounts for the very small latitude errors in Fig. 7. Considering all other system error contributions, an alignment error of 5 arc-sec is quite tolerable and will very likely be attained, at least for the gyro noise input alone.

D. Strapdown Electrostatically-Suspended Gyro Drift Math Model Development,

V. A. Karpenko and D. H. Lipscomb

1. Introduction

A mathematical drift model is being developed for the electrostatically-suspended gyro (ESG) as a part of the Strapdown Electrically-Suspended Gyro Aerospace Navigation (SEAN) system project. The development of the drift math model has been divided into Phase I and Phase II. The Phase I effort was concerned with the development of a stationary drift model based on a postulated theory, and regressed on a sufficiently large and universal set of stationary test data for a single gyro. The main objectives of this phase were as follows:

(1) To develop the basic software system for processing the ESG drift data.

(2) To assess the validity of the present approach to ESG drift modeling.

(3) To develop an initial drift compensation model.

(4) To assess the range of quantitative drift compensation possible with the assumed approach under SEAN operational conditions.

The basic software, as subsequently described, has been thoroughly checked out and has been used successfully in exercising the Phase I math model. The software is sufficiently flexible so that more general math models can be easily inserted into its structure. Positive results were also obtained for the objective listed as (2), above, in that the development of the software, and of the math model, has always progressed smoothly and without any major reversals. The linear model of this phase has been tested under general conditions, and the quantitative results are classified. However, as these tests were conducted only under static conditions, it has been impossible to meet the objective listed as (4), above, to any large degree.

The Phase II effort will concentrate on improving the drift math model, testing the gyro under a wider set of environmental conditions, and testing the developed math model in the SEAN system.

The remainder of this article describes the underlying ideas, basic mathematical formulae, and organization of the computer software developed during the Phase I effort.

2. Equations of ESG Dynamics

The physical makeup of the ESG is described in SPS 37-36, Vol. IV, pp. 51–55. A rotationally-symmetric rotor is suspended in what is assumed to be a potential force field (electrostatic suspension) stationary in the gyro housing. Steady-state conditions prevail so that the following four well-defined centers can be recognized:

(1) The rotor center of mass.

(2) The rotor center of geometry (rotor symmetry center).

(3) The null of the potential force field (electrostatic suspension center).

(4) The rotor electric center.

An orthogonal coordinate frame

$$(i_\alpha) \triangleq \underline{\text{col}}\,(\hat{i}_1, \hat{i}_2, \hat{i}_3)$$

is fixed for the purpose of analysis in the gyro housing with the unit vectors $\hat{i}_\alpha\,(\alpha = 1, 2, 3)$ nominally along the ESG pickoff axes. Another orthogonal frame (S_α) is fixed in the gyro rotor with the unit vector $\hat{S}_1$ lying along the rotor major principal inertia axis. Referring to the diagram of Fig. 8, the following definitions are made:

O = origin of (i_α)—coincides with null of the potential force field

O' = rotor geometry center—origin of (S_α)

O'' = rotor electric center with respect to the potential force field

$\bar{\rho}$ = position of rotor center of mass (CM) in (i_α)

dA = infinitesimal rotor surface area

p = position of dA in (S_α^r)

$(S_\alpha^e)_n$ = Euler frame of the rotor at some reference time t_n

$\hat{N}$ = local unit vector to dA

(e_α) = arbitrary reference inertial frame, which locates the frame (i_α) by position vector $\bar{R}$

The frame (i_α) rotates with angular velocity $\bar{\Omega}$ with respect to (e_α), and the frame (S_α) is rotated by $\bar{\omega}$ with respect to (i_α). The resultant of all applied forces is $\bar{F}$ and $\bar{L}_0$ is their moment about the center O. Then the rotor translational dynamics are described by:

$$\bar{F} = M\left[\ddot{\bar{R}} + \overset{oo}{\bar{\rho}} + 2\bar{\Omega} \times \overset{o}{\bar{\rho}} + \dot{\bar{\Omega}} \times \bar{\rho} + \bar{\Omega} \times (\bar{\Omega} \times \bar{\rho})\right] \tag{1}$$

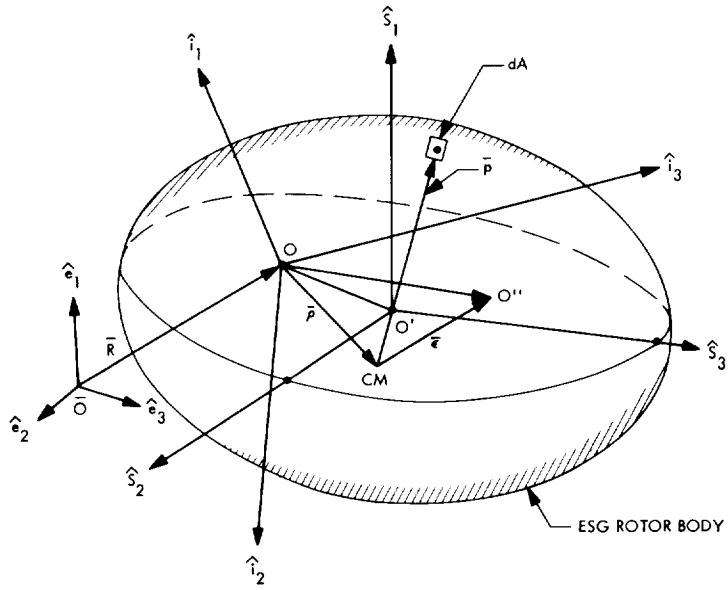


Fig. 8. Schematic of ESG rotor reference frames and significant centers

and rotor rotational dynamics are given by

$$\bar{L}_o = M\{\bar{\rho}\times\ddot{\bar{R}} + \bar{\Omega}\times(\bar{\rho}\times\overset{\circ}{\bar{\rho}}) + \bar{\rho}\times\overset{\circ\circ}{\bar{\rho}} + 2\bar{\rho}\cdot\overset{\circ}{\bar{\rho}}\,\bar{\Omega} - \overset{\circ}{\bar{\rho}}\cdot\bar{\Omega} - \overset{\circ}{\bar{\rho}}\cdot\bar{\Omega} - \bar{\Omega}\times\bar{\rho}\bar{\rho}\cdot\bar{\Omega} + \bar{\rho}^2\dot{\bar{\Omega}} - \bar{\rho}\bar{\rho}\cdot\dot{\bar{\Omega}}\} + \bar{\bar{J}}\cdot(\dot{\bar{\Omega}} + \dot{\bar{\omega}}) + (\bar{\Omega} + \bar{\omega})\times\bar{\bar{J}}\cdot(\bar{\Omega} + \bar{\omega}) \tag{2}$$

where

$\bar{\bar{J}}$ = rotor inertia dyadic,

M = rotor mass,

$\overset{\circ}{\bar{\rho}}$ = time derivative of ρ in frame (i_a), etc

In general, Eqs. (1) and (2) are coupled. In reality, however, the vectors $\overset{\circ}{\bar{\rho}}, \overset{\circ\circ}{\bar{\rho}}$ are, on the average, zero, so that Eq. (2) reduces to equations of the linear gyroscope theory (Ref. 1), and Eqs. (1) and (2) become decoupled as an approximation.

3. The Drift Torque Model

Within the space (e_a), the ESG rotor undergoes motion about the instantaneous center of rotation, which will be referred to here as the kinematic center. Under the expected operational conditions, $|\bar{\omega}| >> \omega_n$ (ω_n is equal to the natural frequency of the gyro suspension and the kinematic center lies approximately on the major principal inertia axis of the rotor, Ref. 2).

The gyroscope will drift about its kinematic center so that analysis, via the linear theory approximation of Eq. (2), requires an appropriate expression of net torque acting on the rotor about this center. The following three generic types of drift torques were considered in the current model:

(1) Torque due to "average" applied gravitational and dynamic accelerations (the torque due to gravitational field curvature and gradient is assumed negligible)

$$\bar{L}_g = M\bar{\varepsilon}\cdot\hat{S}_1\hat{S}_1\times\bar{A} \tag{3}$$

where $\bar{A}$ = net average applied acceleration.

(2) Torque due to magnetic field forces, or similar phenomena

$$\bar{L}_m = -\bar{p}\times\bar{l} \tag{4}$$

where

$\bar{p} = p_a\hat{S}_a$ = magnetic dipole associated with the rotor

$\bar{l} = l_a\hat{i}_a$ = uniform magnetic field flux density associated with the gyro housing

(3) Torque due to electrical field forces

$$\bar{L}'_e = \iint_S \frac{(\bar{p} - \bar{\varepsilon}_a)\times\hat{N}\cdot F_1(v)}{h_a\cdot h_a}\,dA \tag{5}$$

where the integration is performed in the frame (i_a) over the total rotor surface S

and

h_a = actual local rotor–stator gap

$F_1(v)$ = coefficient showing dependence of the electrical torque on the applied voltage v

$\bar{\varepsilon}_a$ = vector defining position of the kinematic center from the average geometry center O'_a of the rotor

4. The Discrete Form of the ESG Drift Equations[2]

Whether in testing or in operational use, the ESG drift is measured (observed) at discrete time instants t_n, where n is the index. It is convenient to define ESG drift as the angular displacement of the gyro spin vector $\hat{S}_1$ in the reference frame $(S'_\alpha)_n \triangleq$ a particular $(S^e_\alpha)\,|\,t_0$; as time grows from t_0. Let $\Delta\theta_n, \Delta\psi_n$ be the drift of $\hat{S}_1$ at time t_n about the $\hat{S}'_{20}$ and $\hat{S}'_{30}$ axes, respectively. Then, it can be shown that the fundamental drift observation equation for the ESG is of the form

$$\begin{bmatrix}\Delta\theta_{n+1}\\ \Delta\psi_{n+1}\end{bmatrix} = \frac{\tau_n}{J_1\omega_n}\begin{bmatrix}-L_3(\alpha_n, EA_0^T, \lambda, t_n - t_0)\\ L_2(\alpha_n, EA_0^T, \lambda, t_n - t_0)\end{bmatrix} + \begin{bmatrix}\Delta\theta_n\\ \Delta\psi_n\end{bmatrix} + \text{noise} \tag{6}$$

where

$J_1\omega_n$ = ESG total angular momentum at t_n

$\tau_n = t_{n+1} - t_n$

[2]V. A. Karpenko has described the complete mathematical development in a series of JPL internal publications.

L_2, L_3 = projections of the ESG drift torque $\Sigma\bar{L} = \bar{L}_g + \bar{L}_m + \bar{L}_e$, [see Eqs. (3), (4) and (5)] on axes $\hat{S}'_{2o}, \hat{S}'_{3o}$, respectively

α_n = direction cosines of $\hat{S}_1$ with respect to (i_α) at time t_n

λ = local latitude

EA_0^T = the definition of orientation of the gravity vector in the reference frame $(S'_\alpha)_0$

Equation (6) is a discrete-form approximation of the linear theory (uncoupled gyrodynamics) valid for small drift angles $\Delta\theta$, $\Delta\psi$, and for a system sampling rate sufficiently high with respect to the ESG housing rotation rate $\bar{\Omega}$. Also, it is assumed that the decay of ω_n is negligible for any normal sampling interval τ_n.

For regression analyses of the drift model, the basic Eq. (6) has been recast into

$$\begin{bmatrix} \Delta\theta_{ni} \\ \Delta\psi_{ni} \end{bmatrix} \triangleq \begin{bmatrix} \Delta\theta_{n+1} \\ \Delta\psi_{n+1} \end{bmatrix} - \begin{bmatrix} \Delta\theta_n \\ \Delta\psi_n \end{bmatrix} = \frac{\tau_n}{J_1\omega_n} [\bar{A}]\,\bar{L} + \text{noise} \tag{7}$$

where

$\bar{L} = n$ − vector of constant regression coefficients

$[\bar{A}] = 2 \times n$ matrix, function of α_n, EA_0^T, λ, and $t_n - t_0$

5. Organization of the Data Processing Program

As a part of the ESG drift model effort, a software system, with an organizational structure as shown in Fig. 9, was conceived with the following two inclusive objectives in mind:

(1) Processing of the ESG drift test data for drift assessment.

(2) Development of the ESG drift mathematical model.

The software system is logically divisible into the following four major subsystems (Fig. 9):

(1) Test data survey, editing, and preparation program (Program A).

(2) The main data processing program (Program B). This program accepts drift data (from the data library) and drift model structure {i.e., the form of $[\bar{A}]$ and $\bar{L}$ in Eq. (7)} as inputs and computes the model regression coefficients $\bar{L}$, the actual gyro drift $\Delta\theta_n, \Delta\psi_n$, and the drift compensation or the drift residuals.

(3) The program for statistical evaluation of the drift residuals using the Monte Carlo approach (Program C). The size of the model, the drift data, and the drift compensation time span are the controllable items.

(4) The drift data simulator program (Program D). This program employs the observation Eq. (6) to compute an artificial (simulated) set of drift data based on assumed concrete drift parameter values and form of the drift torque.

6. Data Processing Activities

The steps involved in processing the raw test data received from the laboratory through to a set of drift coefficients for the ESG are described below.

Gyro drift-test data is received from the laboratory in the form of a paper tape punched on a Flexowriter in the Friden SPS code. This data is converted to punched data cards in a binary-coded decimal (BCD) format through use of a keypunch machine. The data is then operated on by three Honeywell, Inc., generated programs for the IBM 7094 computer to edit out bad data records, add necessary information, and convert data records to the form required for further processing.

The data cards are read into the SM 1100 program that (1) lists the cards, (2) computes direction cosines of the angle between the rotor spin vector and each of the three optical pick-offs, and the rotor speed from the gyro counts read in, (3) edits out data events for several types of possible errors, and (4) writes a magnetic tape in the BCD format. This output tape is read into Program SM 2200 which adds the sign to the computed direction cosine (only two are read at one time), the case-to-earth transformation angles, and the angular momentum. The BCD format is retained by the magnetic output tapes throughout the balance of the processing steps. Noise spikes frequently occur on the drift data and are removed by manually deleting the contaminated data events through use of program SM 3700. These programs have been described in SPS 37-43, Vol. IV, pp. 96–100.

New test data is processed through the Data Survey and Stacking Program to compute the transformation matrices, the gyro drift, and the position with time of the inertial rotor axes in the stator field. This program is also capable of stacking up to seven individual test runs on a composite magnetic tape in the same BCD format. An initial survey of the drift data is desired to assure both the adequate removal of noise spikes and a

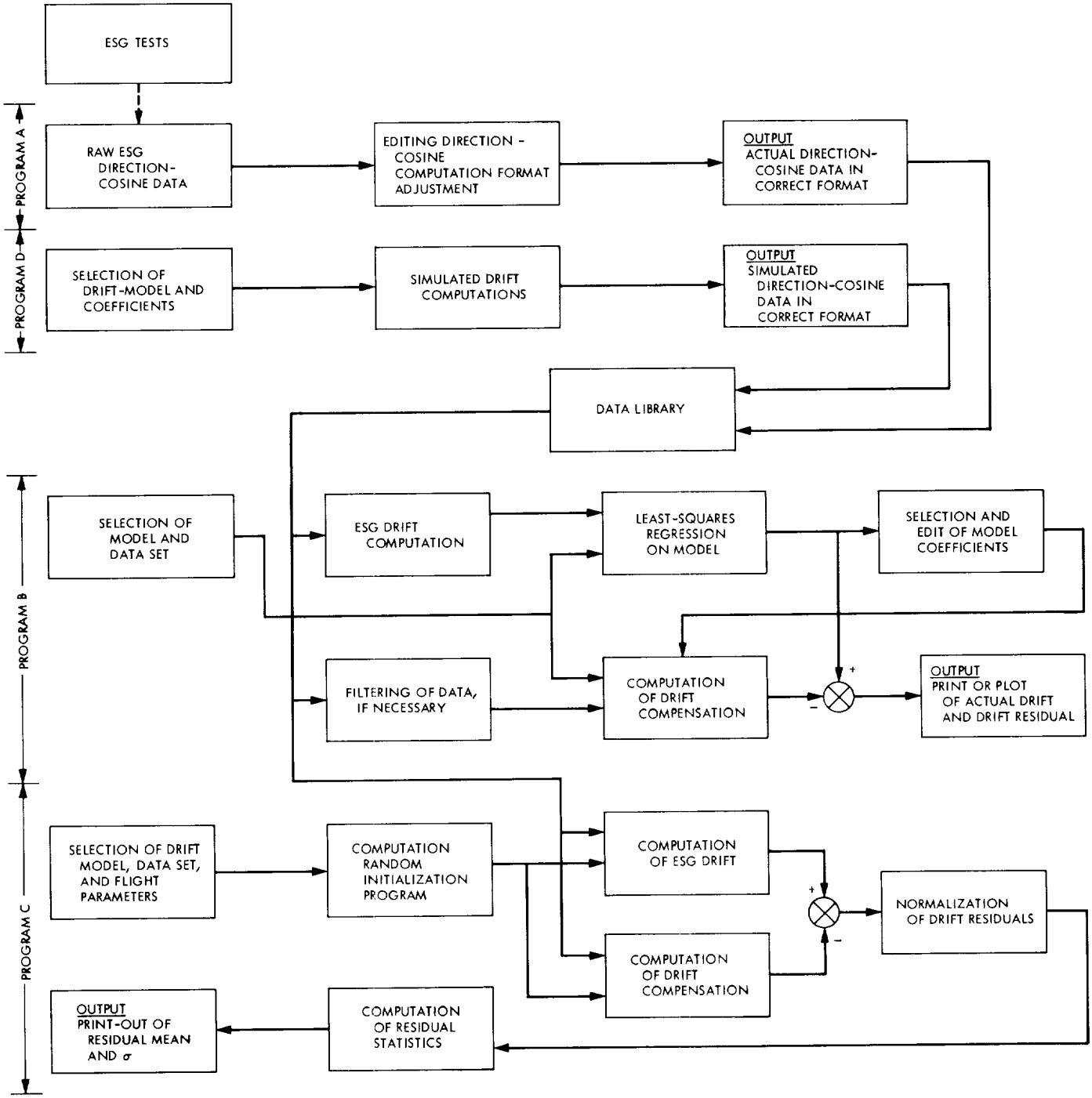
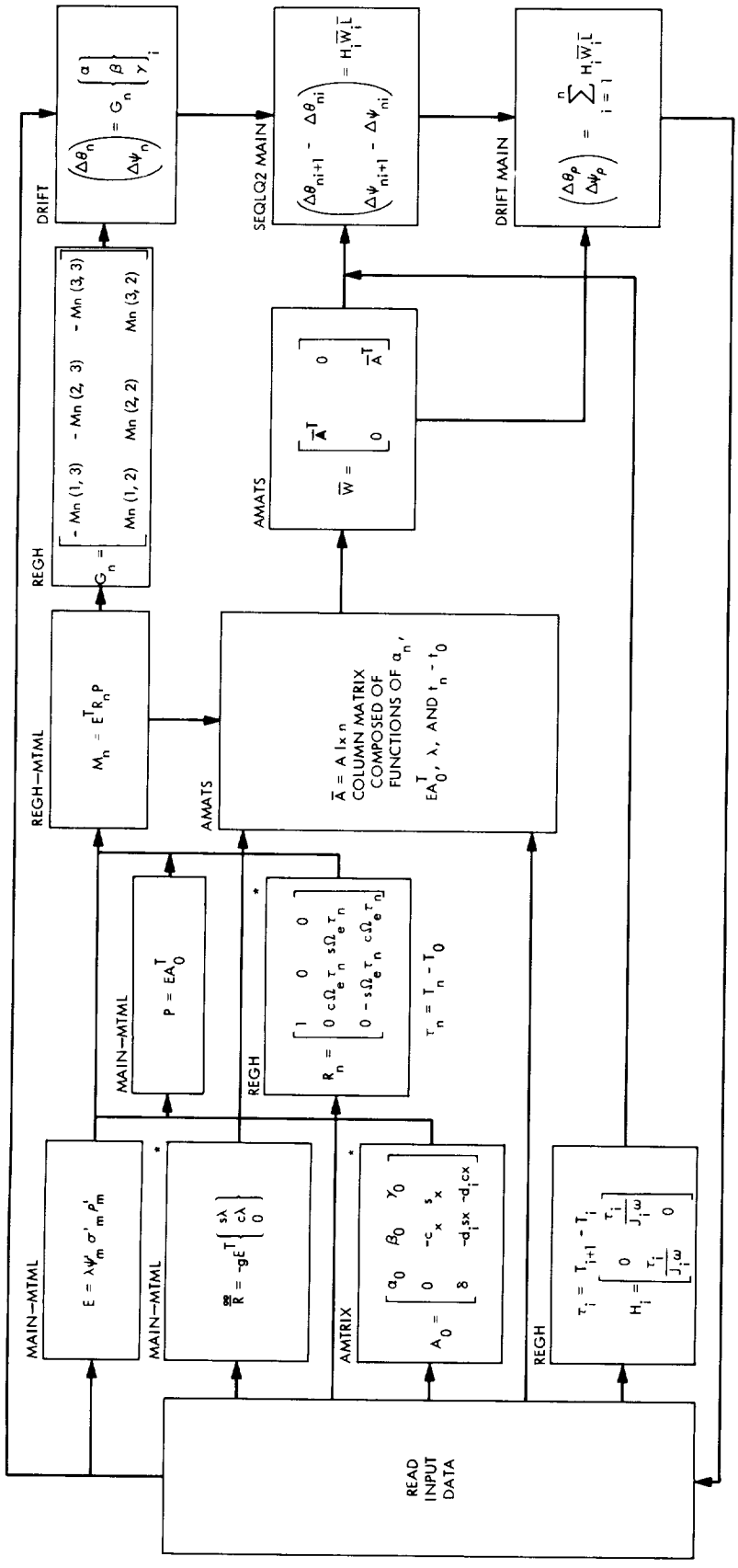

ESG TESTS
PROGRAM A
RAW ESG DIRECTION-COSINE DATA
EDITING DIRECTION - COSINE COMPUTATION FORMAT ADJUSTMENT
OUTPUT ACTUAL DIRECTION-COSINE DATA IN CORRECT FORMAT
PROGRAM D
SELECTION OF DRIFT-MODEL AND COEFFICIENTS
SIMULATED DRIFT COMPUTATIONS
OUTPUT SIMULATED DIRECTION-COSINE DATA IN CORRECT FORMAT
DATA LIBRARY
PROGRAM B
SELECTION OF MODEL AND DATA SET
ESG DRIFT COMPUTATION
LEAST-SQUARES REGRESSION ON MODEL
SELECTION AND EDIT OF MODEL COEFFICIENTS
FILTERING OF DATA, IF NECESSARY
COMPUTATION OF DRIFT COMPENSATION
OUTPUT PRINT OR PLOT OF ACTUAL DRIFT AND DRIFT RESIDUAL
PROGRAM C
SELECTION OF DRIFT MODEL, DATA SET, AND FLIGHT PARAMETERS
COMPUTATION RANDOM INITIALIZATION PROGRAM
COMPUTATION OF ESG DRIFT
NORMALIZATION OF DRIFT RESIDUALS
COMPUTATION OF DRIFT COMPENSATION
OUTPUT PRINT-OUT OF RESIDUAL MEAN AND σ
COMPUTATION OF RESIDUAL STATISTICS

Fig. 9. Conceptual organization of the software system for ESG drift model development


READ INPUT DATA
MAIN—MTML
$E = \lambda \psi'_m \sigma'_m \rho'_m$
MAIN—MTML
$\frac{gg}{R} = -gE^T \begin{Bmatrix} s\lambda \\ c\lambda \\ 0 \end{Bmatrix}$
AMTRIX
$A_0 = \begin{bmatrix} \alpha_0 & \beta_0 & \gamma_0 \\ 0 & -c_x & s_x \\ \delta & -d_i s_x & -d_i c_x \end{bmatrix}$
REGH
$\tau_i = T_{i+1} - T_i$
$H_i = \begin{bmatrix} 0 & \frac{\tau_i}{J_i \omega} \\ \frac{\tau_i}{J_i \omega} & 0 \end{bmatrix}$
MAIN—MTML
$P = EA_0^T$
REGH
$R_n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\Omega_e \tau_n & s\Omega_e \tau_n \\ 0 & -s\Omega_e \tau_n & c\Omega_e \tau_n \end{bmatrix}$
$\tau_n = T_n - T_0$
REGH—MTML
$M_n = E^T R_n P$
AMATS
$\bar{A} = A\ 1 \times n$ COLUMN MATRIX COMPOSED OF FUNCTIONS OF a_n', EA_0^T, λ, AND $t_n - t_0$
REGH
$G_n = \begin{bmatrix} -Mn(1,3) & -Mn(2,3) & -Mn(3,3) \\ Mn(1,2) & Mn(2,2) & Mn(3,2) \end{bmatrix}$
DRIFT
$\begin{pmatrix} \Delta\theta_n \\ \Delta\psi_n \end{pmatrix} = G_n \begin{Bmatrix} \alpha \\ \beta \\ \gamma \end{Bmatrix}_i$
AMATS
$\bar{W} = \begin{bmatrix} \bar{A}^T & 0 \\ 0 & \bar{A}^T \end{bmatrix}$
SEQLQ2 MAIN
$\begin{pmatrix} \Delta\theta_{ni+1} - \Delta\theta_{ni} \\ \Delta\psi_{ni+1} - \Delta\psi_{ni} \end{pmatrix} = H_i \bar{W}_i \bar{L}$
DRIFT MAIN
$\begin{pmatrix} \Delta\theta_p \\ \Delta\psi_p \end{pmatrix} = \sum_{i=1}^{n} H_i \bar{W}_i \bar{L}$

THE NAMES ABOVE EACH BOX IN THE FLOW CHART REFER TO THE SUBPROGRAM NAMES

*s AND c ARE USED TO DENOTE SINE AND COSINE, RESPECTIVELY

Fig. 10. Drift regression and prediction program computational flow chart

reasonable drift magnitude before stacking the new test data with other test data on a composite tape. This composite tape is used to read in the test data for the Drift Regression and Prediction Program. Since only seven input data tapes can be loaded on the IBM 7094 computing equipment, this tape stacking approach greatly increases the variety of input data available for regression and prediction processing on a single computer run. The output plot of the position with time of the rotor axes in the stator field is used to relate drift anomalies with singularities or anomalies of the stator such as pickoff holes and seams which can be manually superimposed on the stator field.

A computational flow chart for the Phase I version of the Drift Regression and Prediction Program is presented in Fig. 10. The coordinate system and the transformations are as follows:

$(\alpha_0, \beta_0, \gamma_0)$ = initial direction cosines of gyro pickoffs with respect to rotor spin vector

R_n = assigns earth rotation

P = transforms rotor inertial-reference frame to inertial earth (earth-centered, polar, equatorial, or local meridian) frame

M_n = transforms rotor inertial frame to current case (pickoff) frame

E = transforms case frame to earth frame

A_0 = transforms initial case frame to rotor inertial frame

The program can be used (1) to derive drift coefficients over selected data segments and predict the drift over these, and other, data segments using the derived coefficients, or (2) to predict the drift over data segments using pre-selected drift coefficients. The data may come from either one or many different laboratory ESG drift tests. There are no additional steps in the program sequence if it is desired to use the data from several drift tests to determine one set of coefficients for the math model.

During the computation of a set of coefficients for the math model, the program first computes the initial earth-to-pickoff, rotor-to-pickoff, and earth-to-rotor transformation matrices, followed by the computation of the actual drift of the earth-fixed gyro. The incremental drift is used in a least-squares regression subroutine to derive the set of drift coefficients, which is then used to compute the predicted drift over the data used in the regression as well as any other test data specified. Correlation matrices are printed out with the coefficients to enable an evaluation of the cross-correlation present. The actual and predicted drifts are plotted together to facilitate an estimation of the gyro compensability. These three outputs (the math model coefficients, their cross-correlation, and the differences between the actual and predicted drift) are used systematically to generate an improved math model and a set of corresponding coefficients. Statistical insight, useful both here and in determining the effect of gyro drift on the SEAN system performance, can be obtained by processing the drift differences in the Statistical Evaluation Program.

References

1. Thomson, W. T., *Introduction to Space Dynamics*, John Wiley & Sons, Inc., New York, 1961.
2. Thomson, W. T., *Vibration Theory and Applications*, Prentice–Hall, Inc., Englewood Cliffs, N. J., 1965.

VII. Guidance and Control Research

GUIDANCE AND CONTROL DIVISION

A. Emitter Work Function of an Operational Converter, *K. Shimada*

1. Introduction

To evaluate the electrical performance of operational thermionic energy converters, it is necessary to determine the emitter work functions. The measurement of such work functions is considerably more difficult in operational converters than in experimental converters because of a lack of accurate knowledge of the emitter areas. Since most of the properties measured, such as current density, are area-dependent, this lack of precision adds uncertainty to the data.

A theory was developed that evaluates the current contributed by the emitter-supporting structure (heat choke) to the total converter current. The theory makes it possible to calculate the correct emitter work functions, since that portion of the current which is truly contributed by the emitter can be evaluated from the total measured current.

The work functions of an emitter of a solar energy thermionic converter were determined by its volt–ampere characteristics in an unignited mode and then applying the theory.

2. Theory

A theory was developed for calculating currents from the side walls of an emitter heat choke for a converter operating in an unignited mode. Since the temperature along the heat choke varies between the emitter temperature T_E and the metal–ceramic seal temperature $T_2 (T_E > T_2)$, the electron emission also varies with the position along the heat choke. This temperature gradient frequently allows the emitter to operate in an ion-rich condition (necessary for work-function measurements), while the cooler end of the heat choke is in an electron-rich condition.

To allow for these varying conditions, the current from the heat choke was calculated as the sum of two currents I_1 and I_2, where I_1 is the current from the ion-rich portion of the heat choke with height X_1 next to the emitter (Fig. 1), and I_2 is the current from the remainder of the heat choke next to the seal.

For a given cesium reservoir temperature, electron *emission* from a cesiated emitter surface of an unignited thermionic diode increases as the emitter temperature decreases. However, the current *conduction* through the diode becomes space-charge limited when the emitter temperature becomes lower than that required for the

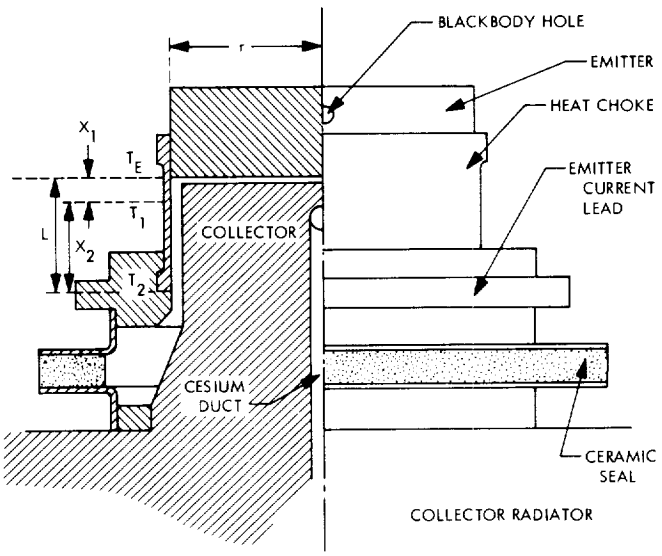


Fig. 1. Emitter region of a thermionic energy converter

neutral emission. Consequently, the current conduction is maximum when the emitter is at a "neutral temperature." Similarly, the largest current is contributed by a small cylindrical area of the heat choke having a temperature approximately equal to the neutral temperature T_1, where the transition from ion-rich to electron-rich emission occurs. This area and, hence, the additional current from the heat choke are calculated as follows.

Let I_1 and I_2 be the currents which are equated to those from annuli with respective heights L_1 and L_2, with a uniform work function ϕ_1 and a uniform temperature T_1. According to the theory,

$$I_1 = 2\pi r L_1 A T_1^2 \exp\left(\frac{-e\phi_1}{kT_1}\right) \tag{1}$$

$$L_1 = X_1 \frac{kT_1}{e\phi_1}\frac{\tau}{\Delta}\left[1 + \frac{2kT_1}{e\phi_1} - \left(1 + \frac{2k\tau}{e\phi_1} + \frac{2\Delta}{T_1}\right)\exp\left(\frac{-e\phi_1\Delta}{kT_1\tau}\right)\right] \tag{2}$$

where

$$\Delta = T_E - T_1 \tag{3}$$

$$\tau = \frac{\Delta}{\left(\frac{\phi_E}{\phi_1} - \frac{T_E}{T_1}\right)} \tag{4}$$

with

r = radius

A = Richardson's constant

e = electron charge

k = Boltzmann's constant

and

$$I_2 = 2\pi r L_2 A T_1^2 \exp\left(\frac{-e\phi_1}{kT_1}\right) \tag{5}$$

$$L_2 = X_2\left(1 - \frac{T_2}{T_1}\right)\frac{e\phi_1}{kT_1} \tag{6}$$

To arrive at the above expression, the temperature distribution along the heat choke was approximated by a linear distribution, and the local work functions were assumed to vary with temperature, according to the Rasor–Warner theory (Ref. 1). Furthermore, the assumptions $\Delta << T_E$ and $e\phi_1/kT_1 >> 1$, which are always true for any conventional thermionic converter, were made in obtaining Eqs. (2) and (6), respectively.

The magnitude of the equivalent length $L_1 + L_2$ turned out to be 20% of the total length of the heat choke in a converter which was tested for the effect of side-wall emission. This increase in emission area was responsible for approximately 60% of the total measured current. Side-wall currents of this magnitude are enough to offset the measured work functions by as much as 0.15 V.

3. Emitter Work Functions of SN 107

Emitter work functions were determined for a hardware-type thermionic converter,[1] serial number SN 107. The electrode geometry (Fig. 1) was plane-parallel with an interelectrode gap of 0.0045 in. (0.0114 cm). A cylindrical heat choke (r = 0.8 cm, wall thickness = 0.0076 cm) separated the hot emitter from the cold metal–ceramic seal. In the analysis of its side-wall emission, the effective height L of the heat choke was assumed to be 0.5 cm. The emitter, emitter heat choke, and the collector were rhenium, with an assumed uncesiated work function of 4.8 eV (Ref. 2). To assess the effect of side-wall currents, the cesiated work functions were first calculated from the measured saturated currents I_m of the converter, and then from the true emitter current $(I_m - I_1 - I_2)$ of the converter.

[1]Fabricated to JPL specifications by Electro-Optical Systems, Pasadena, Calif.

Procedures for calculating the apparent and the true emitter work functions of the converter SN 107 are:

(1) For *apparent work functions,* use the Richardson equation with the A-value of 120 A/cm^2-°K^2. To determine the saturation-current density, divide the measured saturation currents by the geometrical emitter area of 2 cm^2. The heat-choke area is not considered in any way in this type of determination.

(2) For *true work functions,* determine the emitter temperature T_E, the cesium reservoir temperature T_{Cs}, and the seal temperature T_2. Assume that the electrode materials are clean, and hence the uncesiated work functions are 4.8 V, and that the cesiated work functions are governed by the Rasor–Warner theory. Determine the neutral emission parameters T_1 and ϕ_1 from the Rasor–Warner theory. Establish X_1 by assuming a linear temperature distribution along the heat choke ($L = 0.5$). Calculate I_1 and I_2 from Eqs. (1) and (5). Subtract $I_1 + I_2$ from the measured saturation current I_m. Apply the Richardson equation to determine the work functions. Average this calculated work function with the theoretical value for cesiated rhenium to arrive at the true work function.

Results are shown in Fig. 2. Here, the T_E/T_{Cs} was corrected for the transpiration effect (Ref. 3) of cesium gas, especially for low cesium temperatures T_{Cs} for which the mean-free-path of the cesium atoms was equal to or larger than the interelectrode gap. Comparison of these results showed that measured work functions were consistently lower, by as much as 0.15 V, than those which were expected for a rhenium emitter. On the other hand, the work functions, which were determined from the true emitter current using the theory, agreed with expected values within 0.02 V.

4. Conclusion

A conduction current as large as 60% of the measured current could originate from the side wall of a heat choke

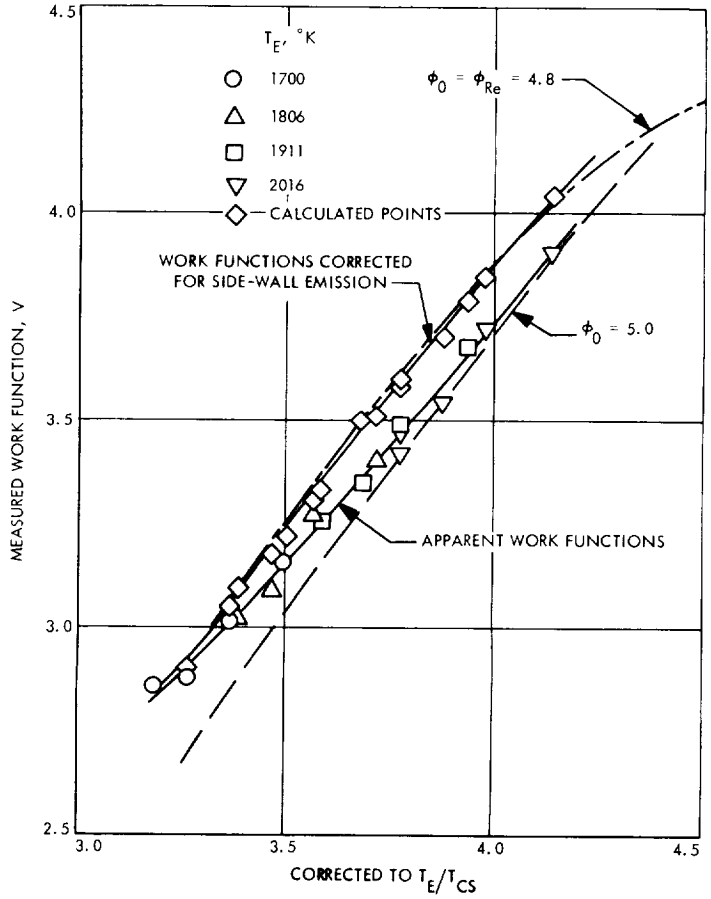


Fig. 2. Comparisons between apparent and corrected emitter work functions for SN 107

in an operational thermionic energy converter when it is operated in an ion-rich, unignited mode. Such currents are equivalent to those that are contributed by approximately 20% of the side-wall area operating under neutral-emission conditions. The side-wall current must be subtracted from the total measured current to obtain meaningful values for the emitter work function in operational converters; otherwise, errors as large as 0.15 V would result in work functions.

An application of the theory for calculating side-wall current and the true emitter work function of a hardware converter SN 107 demonstrated that the true work functions were indeed within 0.02 V of the theoretical values.

References

1. Rasor, N. S., and Warner, C., III, "Correlation of Electron, Ion and Atom Emission Energies," *First Summary Report of Basic Research in Thermionic Energy Conversion Processes*, Report No. AI-6799, pp. 45–79. Atomics International Division, North American Aviation, Inc., Canoga Park, Calif., Nov. 15, 1961.
2. *Annual Technical Summary Report*, p. V-20. Air Force Contract AF 19(604)-8453, No. TE 7-65, Thermo Electron Engineering Corp., Waltham, Mass., 1965.
3. Kennard, E. H., *Kinetic Theory of Gases*, pp. 66–67. McGraw-Hill Book Co., Inc., New York, 1938.

B. Surface Barriers on Layer Semiconductors: GaSe,[2] S. Kurtin[3] and C. A. Mead[3]

1. Introduction

Metallic surface barriers have been investigated on a wide variety of materials by many researchers. On the basis of information gained from their research, semiconductor materials have been divided into two broad classes (Ref. 1). The first class is composed of materials such as ZnS in which the barrier energy depends directly on the electronegativity of the metal on its surface. Such behavior is interpreted to mean that ZnS has a low density of surface states ($<10^{12}$ eV^{-1}-cm^{-2}). The second class is composed of materials such as GaAs in which the barrier energy is nearly independent of the metal on its surface. This type of behavior is interpreted to mean that GaAs has a high density of surface states ($\simeq 10^{14}$ eV^{-1}-cm^{-2}) with the Fermi level being set at approximately $E_g/3$, where E_g is the energy gap. Ionic solids fall into the first class whereas covalent solids, with their "dangling bonds," fall into the second.

Layer compounds possess a high degree of anisotropy in their chemical bonding (Refs. 2 and 3) and physical properties (Ref. 4). These compounds are of interest because of this anisotropy and because study of their surface barriers may aid in understanding the relationship between chemical bonding and surface properties. In this study, the behavior of surface barriers on one layer compound, GaSe, was investigated by the photoresponse method.

2. Sample Preparation

Samples were prepared by peeling small ($\sim 8 \times \sim 3$ mm) single crystal flakes from a Bridgman-grown boule of p-type ($p \simeq 8 \times 10^{13}$) GaSe. An ohmic contact was made to one side of each flake by vacuum evaporation of a thin (~ 50 Å) platinum layer followed by a thick (~ 1000 Å) zinc layer and subsequent alloying in a hydrogen-reducing atmosphere at approximately 450°C for 1 min. The platinum layer is necessary to provide nucleation sites on the smooth, inert GaSe surface. Without such a layer, zinc atoms will run along the surface until they encounter a cleavage step or edge and a uniform layer cannot be obtained. Electrical contact was made to the zinc by low-temperature soldering with In–Ag solder.

Different methods were used to construct surface barriers for each metal. For the more noble metals, the procedure consisted of cleaving the GaSe sample in air to expose a fresh (0001) surface and then vacuum depositing, through a wire screen mesh, small ($\sim 0.010 \times 0.010$ in.) semitransparent metallic dots. Samples of this type were well-behaved for periods of minutes to weeks depending on the metal used.

The above technique was not directly applicable to the alkali and alkali earth metals since they oxidize rapidly on exposure to air, being especially vulnerable in thin film form. It is conceivable to make a two-layer vacuum deposition with a relatively inert metal covering an active one. The registration problems involved in such a procedure are severe; if any of the covering metal overlaps the active metal dot, the measured barrier potential will be primarily that of the more noble metal since barrier height decreases with increasing electronegativity for metals on p-type semiconductors. An interesting variation of this technique was devised. It consists of covering the entire GaSe surface with a transparent layer of active

[2]The authors acknowledge the assistance of J. L. Gumnick and C. D. Hollish of ITT Laboratories and H. M. Simpson, L. G. Fishbone, and R. S. Douglass of Caltech.

[3]California Institute of Technology, Pasadena, Calif.

metal and then evaporating, through a mesh, semitransparent dots of noble metal. This entire procedure is carried out at 2×10^{-7} torr in approximately 30 s. When a completed sample is removed from the vacuum deposition apparatus, all the dots are found to be shorted together. As the active metal oxidizes, the region between dots can no longer conduct and photoresponse measurements may be made. This approach allows measurements to be made by a straightforward front wall cell technique in atmospheric ambient environments.

Metals used for barriers were evaporated from directly heated tungsten filaments with the exception of cesium, for which a Cs_2CrO_4-Si channel was used.

3. Photoresponse Measurements

Barrier energies in the GaSe surface barrier structures were measured primarily by the photoresponse method. Photocurrents were measured with a 50-Hz chopped-light system and lock-in detector. The light source was a Sylvania sun gun lamp with a quartz prism monochrometer. Intensity calibration was done with a Reeder thermocouple.

The dependence of photocurrent on photon energy was generally found to follow a square law as expected from the simple theory of photoemission from metals.

Typical curves of the square root of photoresponse per incident photon versus photon energy for several metals appear in Fig. 3. Intercepts on the photon energy axis correspond to the GaSe–metal barrier energy measured from the GaSe valence band.

From photoresponse data, as in Fig. 3, it is possible to plot barrier energy versus electronegativity for the different metals. Such a curve appears in Fig. 4, where the

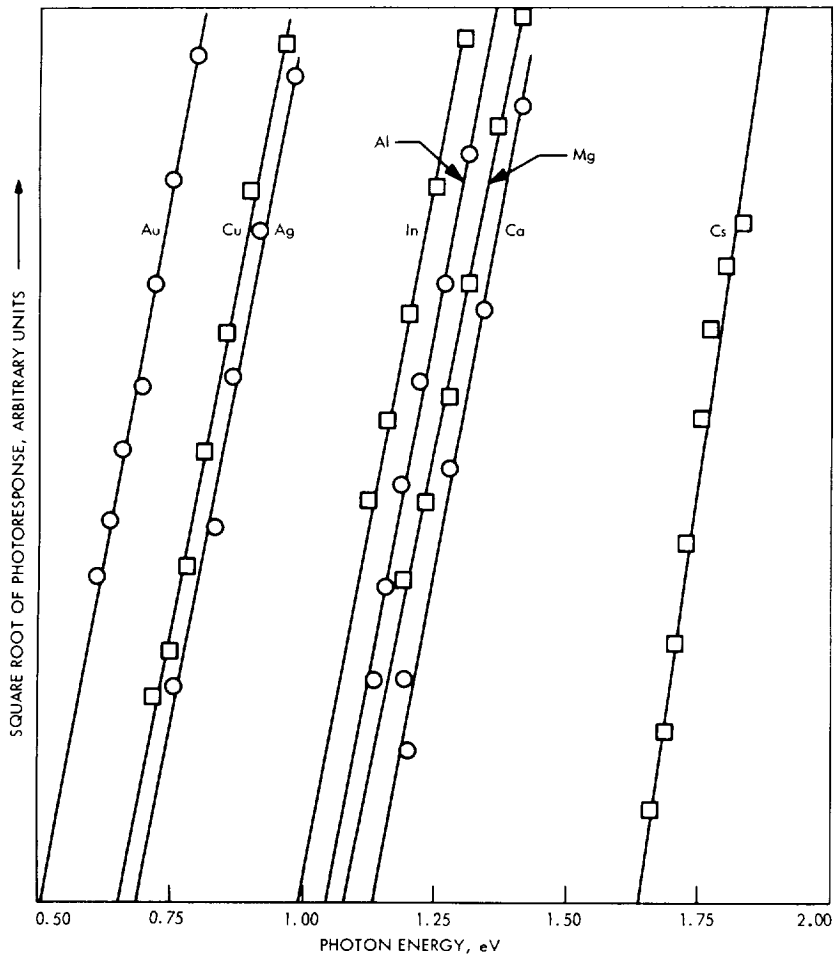


Fig. 3. Typical plots of photoresponse per incident photon vs photon energy for several metals

values are the best averages of several samples. The slope of the reference line is $\simeq 0.6$.

4. Conclusion

The experimental results as presented in Fig. 4 differ in character from those observed in nonlayer compounds.

As discussed, ionic materials are characterized by unity slope, and covalent materials by a slope of approximately 0.1 on such a plot. A result of slope $\simeq 0.6$ may be interpreted (Ref. 5) in terms of a uniform density of surface states ($\simeq 7 \times 10^{12}$ eV^{-1}-cm^{-2}) distributed over the forbidden gap. This behavior is evident over a range of electronegativity from 0.7 to 2.5 V, thus giving some credence to the linear relationship proposed.

From the physical nature of layer compounds, it is evident that their chemical bonding is fundamentally different from that of typical crystalline semiconductors. The unusual dependence of surface barrier potential on electronegativity that was observed may be intrinsic to the layer structure. If so, a deeper understanding of the relationship between chemical bonding and surface properties may be gained through the study of layer compounds.

References

1. Mead, C. A., *Solid State Electronics*, Vol. 9, p. 1023, 1966.
2. Basinski, Z. S., Dove, D. B., and Mooser, E., *CR Acad. Sci.*, Paris, Vol. 34, p. 373, 1961.
3. Fisher, G., and Brebner, J. L., *J. Phys. Chem. Solids*, Vol. 23, p. 1363, 1962.
4. Leung, P. C., et al., *J. Phys. Chem. Solids*, Vol. 27, p. 849, 1966.
5. Cowley, A. M., and Sze, S. M., *J. Appl. Phys.*, Vol. 36, p. 3213, 1965.

C. Noise Measurements on a Double-Injection Silicon Diode, *D. H. Lee,[4] and H. R. Bilger,[5] and M-A. Nicolet[4]*

1. Introduction

Nicolet, Bilger and McCarter (Ref. 1) report noise measurements on a double-injection silicon diode (DISD) in which the current–voltage (I–V) characteristic has a linear and quadratic range. They demonstrate that the white noise level can be represented by a noise current generator $\langle i^2 \rangle^{1/2}$, where

$$\langle i^2 \rangle = \beta \cdot 4kT\,(\partial I/\partial V)\,\Delta f \qquad (1)$$

[4]California Institute of Technology, Pasadena, Calif.

[5]Oklahoma State University, Stillwater, Okla.

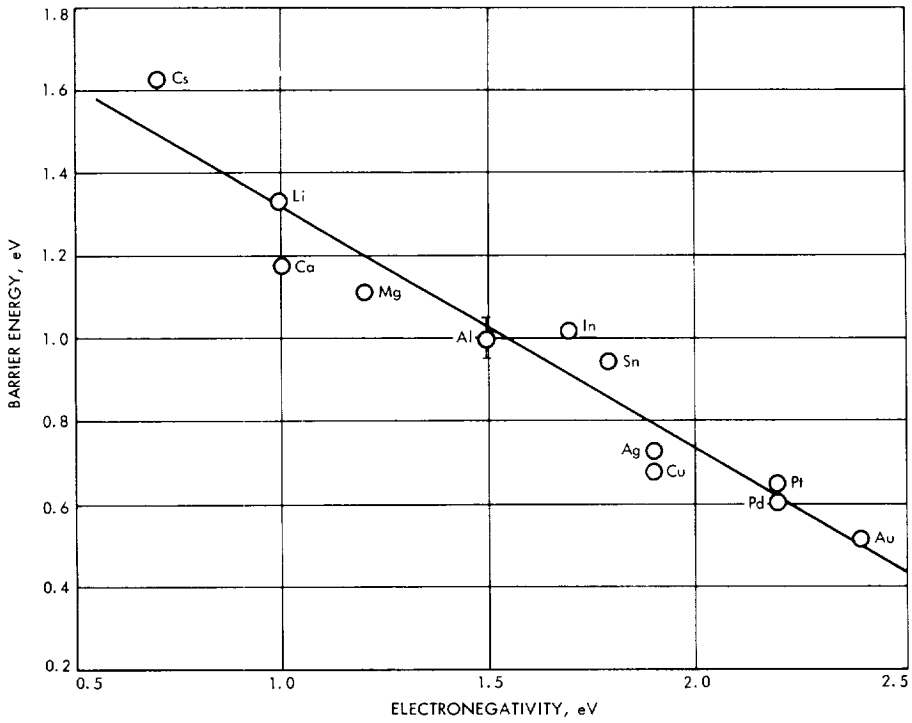


Fig. 4. Barrier energy vs electronegativity of metal

and

k = Boltzmann's constant

T = temperature

$\partial I/\partial V$ = low-frequency conductance of the device

Δf = frequency range

The factor $\beta = 1$ holds for the linear portion of the I–V characteristic and $\beta = 1/2$ for the quadratic range. In similar measurements of the noise in a germanium double-injection diode, Liu, Yamamoto, and van der Ziel (Ref. 2) indicate that the white noise level is represented by a noise current generator $\langle i^2\rangle^{1/2}$, where

$$\langle i^2\rangle = \alpha \cdot 4kT\, g_{hf}\, \Delta f \tag{2}$$

and g_{hf} is the high-frequency differential conductance of the diode. Liu, et al., find $\alpha \approx 1$ for both the linear and quadratic range of the I–V characteristic. Similar results are also obtained by Driedonks, Zijlstra, and Alkemade (Ref. 3) for a germanium double-injection diode which has an $I \sim V^3$ characteristic.

2. Noise and Conductance

Measurements were made at room temperature (298°K) on a DISD identical to that of Ref. 1. The I–V characteristic of the 3.1- × 3.2- × 6.0-mm diode with doping level $\sim 1.1 \times 10^{12}$ cm^{-3} is given in Fig. 5. By comparing the noise current $\langle i^2\rangle^{1/2}$ of the DISD with a temperature-limited calibrator shot noise source (5722 vacuum tube), $\langle i^2\rangle$ can be represented in terms of an equivalent saturated noise current I_{equiv}:

$$\langle i^2\rangle = 2q\, I_{equiv}\, \Delta f \tag{3}$$

where q is the electron charge. Figure 6 shows the noise spectra of the DISD. The continuous curves represent a least squares fit of the function

$$I_{equiv} = A + (B/f^{\gamma}) \tag{4}$$

to the data. The excellent fit of Eq. (4) to the noise data indicates that the noise consists of a frequency-independent (white) component and a "$1/f$" component.

Measurements of the diode differential conductance $g(\omega)$ and the appropriate least squares fit were made from

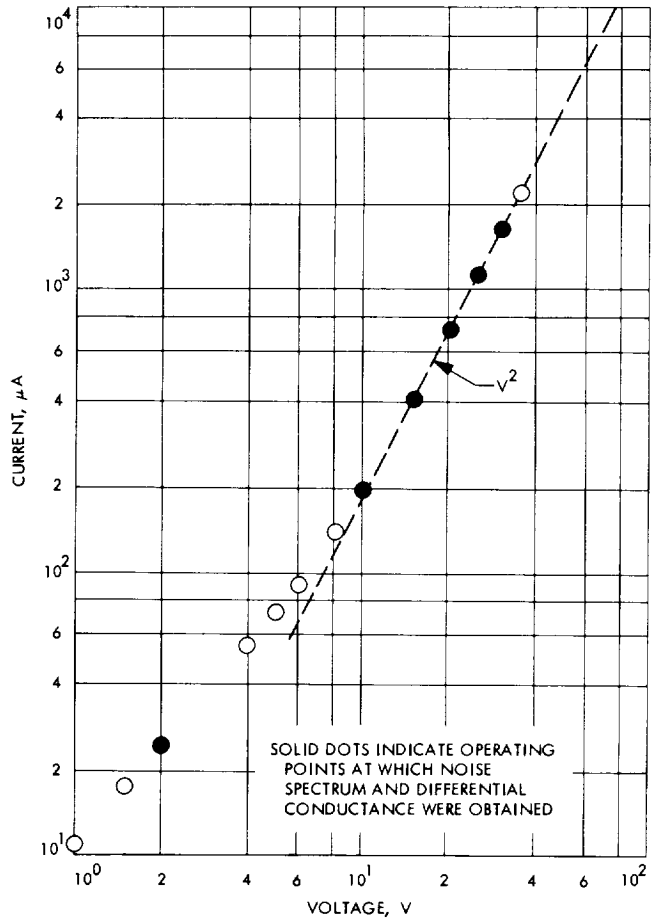

CURRENT, μA
VOLTAGE, V
V²
SOLID DOTS INDICATE OPERATING POINTS AT WHICH NOISE SPECTRUM AND DIFFERENTIAL CONDUCTANCE WERE OBTAINED

Fig. 5. I–V characteristic of 3.1- × 3.2- × 6.0-mm DISD with doping level $\sim 1.1 \times 10^{12}$ cm^{-3} at room temperature

90 Hz to 22 MHz. The result is given in Fig. 7. In the range $I \sim V$, the low-frequency conductance g_{lf} is equal to the high-frequency conductance g_{hf}; whereas, in the $I \sim V^2$ range, $g_{lf} = 2g_{hf}$. The high-frequency conductance is reached when $\omega\tau_{eff} >> 1$, where τ_{eff} is the effective recombination time for the electrons and holes. This time is approximately 38 μs in the present structure and is determined from the curves of least squares fit in Fig. 7. By applying a differential voltage step to the DISD and measuring the current response, a second independent measurement of τ_{eff} was obtained and found to be approximately 37 μs.

3. Discussion

To compare the white noise of the DISD with the Nyquist noise of the high-frequency conductance, one may write

$$\langle i^2\rangle_{DISD} = \alpha \cdot 4kT\, g_{hf}\, \Delta f \tag{5}$$

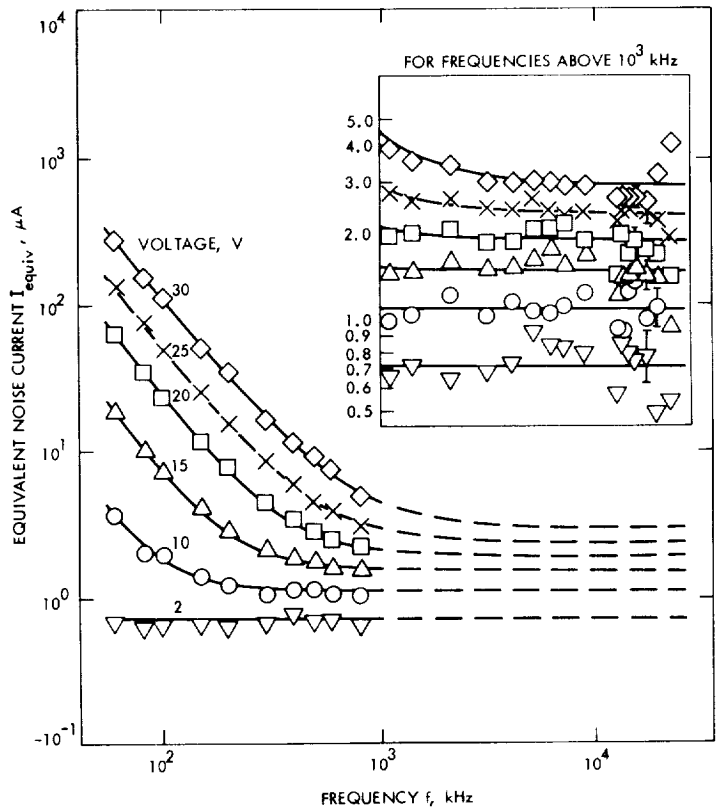


Fig. 6. Noise spectra of the DISD

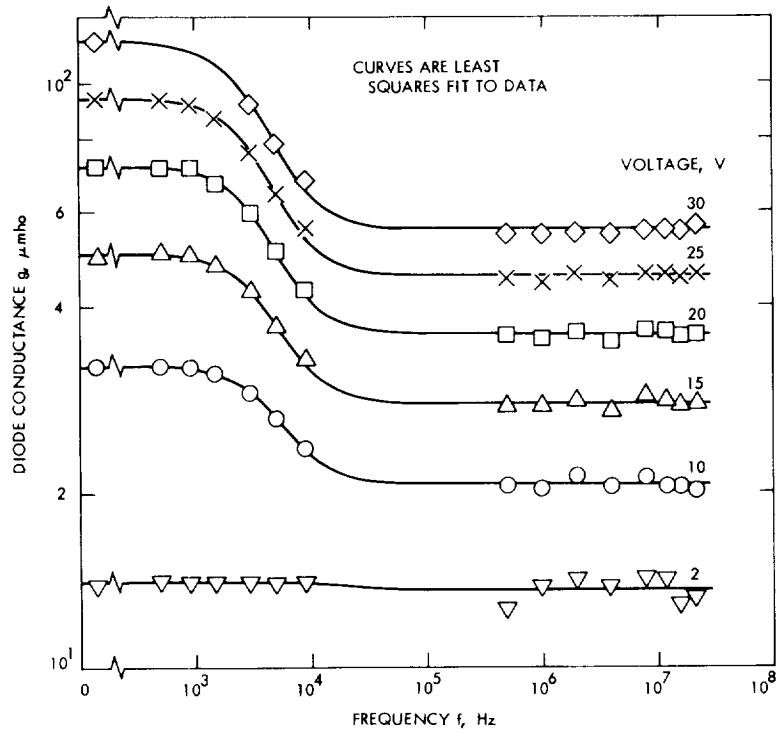


Fig. 7. Diode conductance vs frequency

With Eqs. (3–5), this can be written as

$$A = \alpha \cdot \frac{2kT}{q} g_{hf} \tag{6}$$

The least squares fitted values of A and g_{hf} are plotted in Fig. 8 along with the theoretical value of Eq. (6) for $\alpha = 1.00$. A value of $\alpha = 1.04 \pm 0.05$ is obtained from a least squares fit to the experimental data. The excellent agreement indicates that, for frequencies much greater than $1/\tau_{eff}$, the noise of the device is represented by

$$\langle i^2 \rangle = 4kT\, g_{hf}\, \Delta f \tag{7}$$

This result is compatible with those of Refs. 1–3. It should be noted, however, that the diodes investigated in Refs. 2 and 3 are germanium devices and that their conductances at high frequencies are not constant, as is the case here. The thermal noise of the double injection diode as expressed by Eq. (7) can thus be considered as a general property of the device. This interpretation is evident if one realizes that for frequencies at which the noise current fluctuations are not influenced by generation recombination, the electrons and holes are independently at thermal equilibrium with the lattice, and thus exhibit thermal noise.

A more detailed analysis of the $1/f$ region, where $I_{equiv} \sim B/f^{\gamma}$, yields values of $B = I^{2.2}$ and $\gamma = 2.0$. It is speculated, therefore, that the origin of the $1/f$ noise may be generation recombination.

References

1. Nicolet, M-A., Bilger, H. R., and McCarter, E. R., *Appl. Phys. Letters*, Vol. 9, p. 434, 1966.
2. Liu, S. T., Yamamoto, S., and van der Ziel, A., *Appl. Phys. Letters*, Vol. 10, p. 308, 1967.
3. Driedonks, F., Zijlstra, R. J. J., and Alkemade, C. Th., *Appl. Phys. Letters*, Vol. 11, p. 318, 1967.

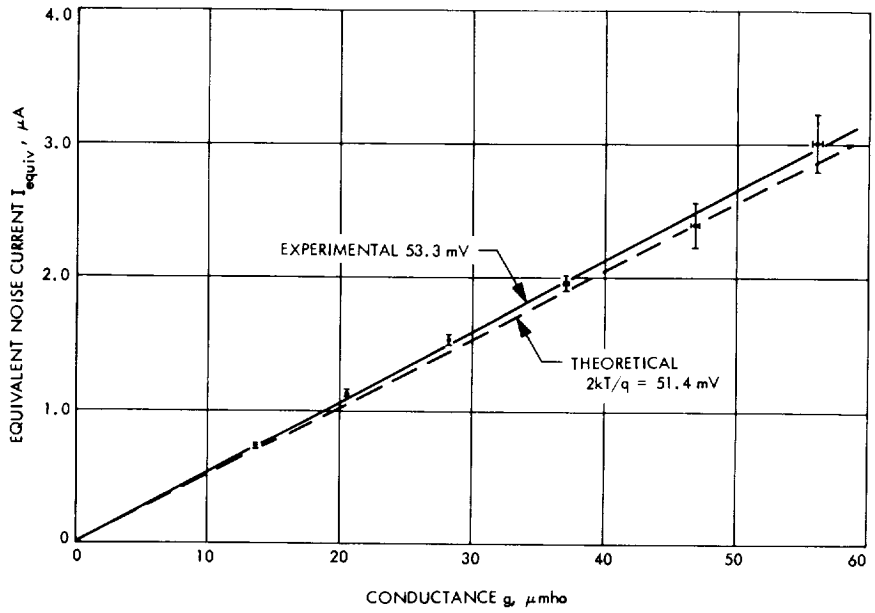


Fig. 8. Equivalent noise current vs conductance at high-frequency levels

VIII. Electronic Packaging and Cabling

ENGINEERING MECHANICS DIVISION

A. Thermal Resistance of Transistors in JEDEC TO-5 and TO-18 Packages, *R. M. Jorgensen*

1. Introduction

In conventional ground environment electronic packaging, the heat developed in electronic components is largely dissipated into the surrounding air by convection. In the space environment, air is lacking, and thus the convection mode of heat transfer is absent. A design technique has been evolved over the years that recognizes the need for a conduction path for heat transfer from transistors packaged in "cans" when printed circuits or terminal boards are part of the packaging concept. This technique places the can of the transistor against a thin epoxy-glass laminate electrical insulating sheet adjacent to a metal heat sink. The method used to quantitatively evaluate the thermal resistance of canned transistors when mounted with the above-mentioned installation technique and the results are described in this article. The standard JEDEC TO-5 and TO-18 transistor packages were chosen for the first evaluation.

2. Mathematical Model

A mathematical model was generated in the form of a thermal resistance network, with each term of the network representing a series conduction path from the semiconductor chip to the heat-sink chassis. Part of the intent of this mathematical model was to identify which portion of the transistor installation provides the limiting factor in increased heat transfer from the transistor. Radiation internal to the transistor can was ignored, due to the complexity of the mathematics.

A generalized model of a mounted transistor is presented in Fig. 1. It is assumed that the header contacts

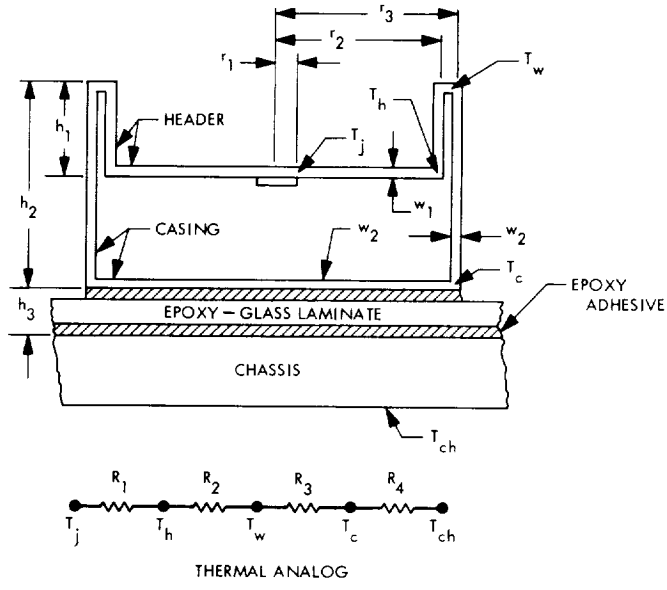


Fig. 1. Cross section of transistor mounted for heat transfer by conduction

the case only at the welded joint. This may not always be true, but it serves as a conservative assumption. The conductivities of the epoxy adhesive and epoxy glass laminate are assumed to be equal, which allows the thickness of adhesive plus insulator to be lumped as h_3. K_h, K_c and K_r are the thermal conductivities of the header, case, and epoxy insulator and bonding material, respectively. T_j, T_h, T_w, T_c and T_{ch} are temperatures at the points indicated on Fig. 1.

R_1, R_2, R_3 and R_4 are expressed in terms of geometry and conductivities as follows:

$$R_1 = \frac{\ln (r_2/r_1)}{2\pi w_1 K_h} \tag{1}$$

$$R_2 = \frac{h_1}{2\pi r_2 w_1 K_h} \tag{2}$$

$$R_3 = \frac{h_2}{2\pi r_3 w_2 K_c} \tag{3}$$

$$R_4 = \frac{Nh_3}{2\pi K_r r_3} \cdot \frac{I_0 (Nr_3)}{I_1 (Nr_3)} \tag{4}$$

where I_0, I_1 are modified Bessel functions and

$$N = \left(\frac{1}{h_3 w_2} \cdot \frac{K_r}{K_c}\right)^{1/2}$$

R_4 was derived by developing an expression for the temperature distribution in a disk, into which heat flows through the perimeter and out through one face. The other face is thermally insulated. The disk is assumed to be thin enough that temperature t can be assumed to be a function of radius r only. The perimeter has uniform temperature T_c. The thermal conductivity of the disk is K_c. The noninsulated face is in perfect thermal contact with a resin bonding compound of thermal conductivity K_r and of thickness h_3.

It can be shown (Ref. 1, p. 82) that the equation for the temperature distribution in a disk (the lid of the transistor case) is

$$t - T_{ch} = \frac{T_c - T_{ch}}{I_0 (Nr_3)} I_0 (Nr) \tag{5}$$

where

I_0 = modified Bessel function

$$N = \left(\frac{1}{h_3 w_2} \cdot \frac{K_r}{K_h}\right)^{1/2}$$

This temperature distribution was used to integrate the heat flow over the entire lid of the transistor case, which is given by

$$q = \int_0^{r_3} 2\pi K_r \frac{(t - T_{ch})}{h_3} r\, dr \tag{6}$$

Substituting Eq. (5) into Eq. (6) and carrying out the integration gives

$$q = \frac{2\pi K_r r_3}{Nh_3} \cdot \frac{I_1 (Nr_3)}{I_0 (Nr_3)} (T_c - T_{ch}) \tag{7}$$

Since

$$R_4 = \frac{T_c - T_{ch}}{q} \tag{8}$$

Eq. (4) is derived by simple substitution of Eq. (7) into Eq. (8).

Several TO-5 and TO-18 transistor cans were measured, and the various pertinent dimensions for computing numerical values of R_1 through R_4 are listed in Table 1. In all cases, the measurement tending to maximize the value of thermal resistance is listed in the table. Values of K were chosen from handbooks. K_c and K_h were chosen to be equal, which provides a further conservative assumption, since K_c may be greater than K_h. (The case is sometimes made of nickel, which is more conductive than the Kovar header by almost an order of magnitude.)

Table 1. Dimensions and conductivities for TO-5 and TO-18 transistor case installations

TO-5			TO-18		
Area	Inches	Centimeters	Area	Inches	Centimeters
w_1	0.014	0.0356	w_1	0.008	0.0203
w_2	0.010	0.0254	w_2	0.010	0.0254
h_1	0.097	0.2460	h_1	0.100	0.254
h_2	0.245	0.6240	h_2	0.200	0.508
h_3	0.018	0.0458	h_3	0.018	0.0457
r_1	0.017	0.0433	r_1	0.017	0.0432
r_2	0.145	0.369	r_2	0.075	0.190
r_3	0.160	0.407	r_3	0.090	0.228

$K_h = 4.6 \times 10^{-2}$ cal/s-cm-°C
$K_c = 4.6 \times 10^{-2}$ cal/s-cm-°C
$K_r = 4.5 \times 10^{-4}$ cal/s-cm-°C

The values of thermal resistance resulting from use of Table 1 data in Eqs. (1) through (4) are given in Table 2.

Table 2. Calculated thermal resistances of TO-5 and TO-18 transistor cans

TO-5	TO-18
R_1 = 208 °C-s/cal	R_1 = 202 °C-s/cal
R_2 = 91	R_2 = 182
R_3 = 183	R_3 = 380
R_4 = 228	R_4 = 667
R_t = 710 °C-s/cal	R_t = 1432 °C-s/cal
= 173 °C/W	= 348 °C/W
= 312 °F/W	= 626 °F/W
= 0.31 °F/mW	= 0.63 °F/mW
$R_t = R_1 + R_2 + R_3 + R_4$.	

3. Experimental Measurements

To determine the heat dissipation capability of various transistors in chassis-mounted circuit boards, tests were undertaken to measure thermal resistance between the junction and the chassis. The relation between thermal resistance R_t, heat dissipation q, and junction and chassis temperatures T_j and T_{ch}, respectively, is expressed as

$$q = \frac{T_j - T_{ch}}{R_t}$$

Thermal resistance R_t can be determined for a given mounted transistor from the above formula if q, T_j, and T_{ch} can be measured while the transistor is operating in the steady state. Measurement of dissipation and chassis temperature is straightforward. Junction temperature, however, cannot be measured directly, for it is not practical to attach a thermocouple to the silicon chip which constitutes the junction.

Several methods of determining junction temperature indirectly, by measuring such temperature-sensitive parameters as the collector-to-base current I_{cbo} or the collector-to-base voltage V_{cbo}, exist. Germanium devices have a temperature-sensitive I_{cbo}, whereas silicon devices have a temperature-sensitive V_{cbo}. The test circuit for using the V_{cbo} technique, which gave significant results, is shown in Fig. 2.

a. Calibration of transistor under test. For each transistor tested, the V_{cbo} (across terminals A and B in Fig. 2)

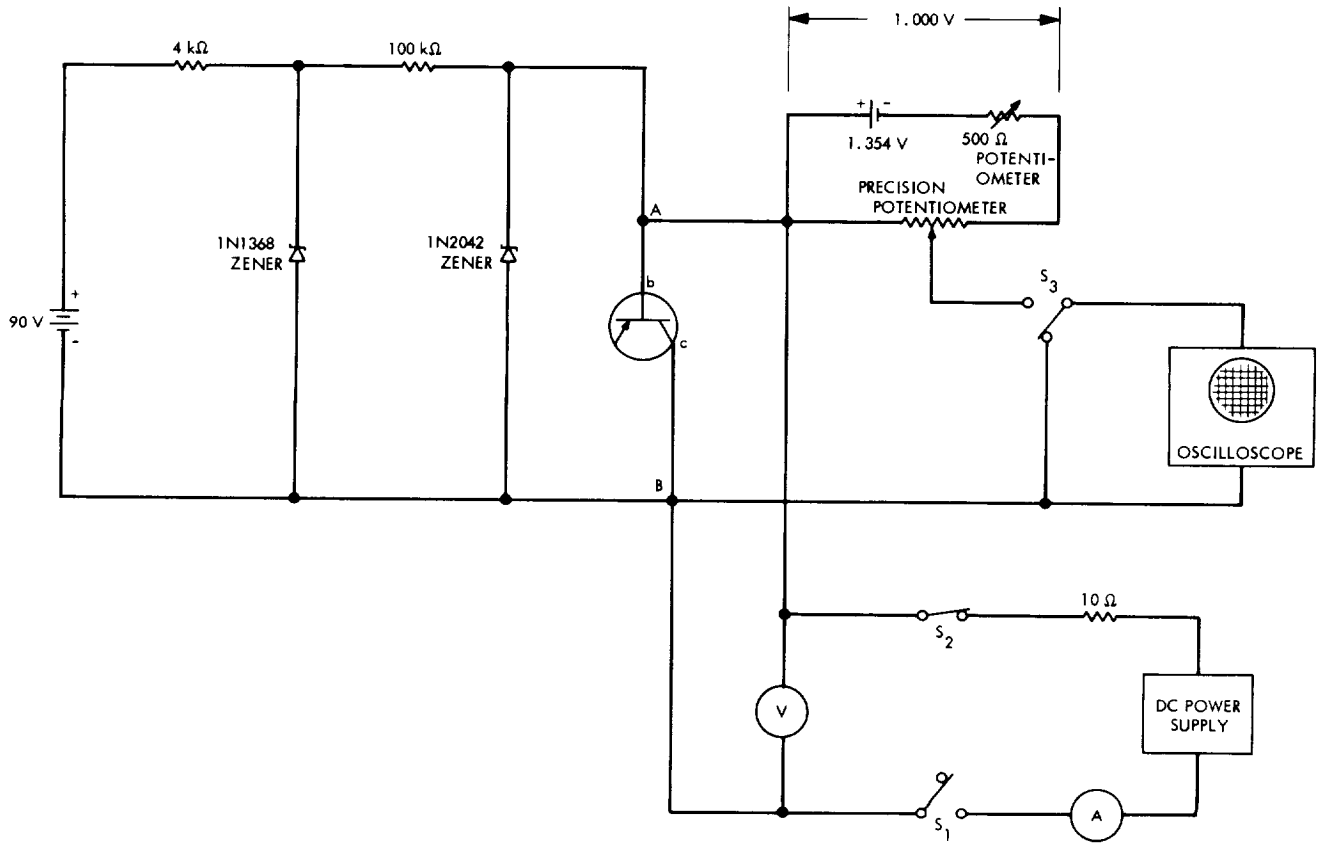


Fig. 2. Test circuit for determining temperature of transistor under power-dissipating conditions

must be calibrated versus junction temperature. For calibration, the load circuit is disconnected by opening switch S1. The chassis containing the mounted transistor is placed in an oven. When the temperature of the oven, the chassis, and the transistor stabilizes, it can be assumed that the junction has the same temperature as the ambient in the oven, since no power is being supplied to the transistor. The transistor is then connected to the circuit at terminals A and B by leads which are brought out through the oven wall. Power is now being supplied to the transistor from the 90-V battery. The Zener diodes, however, limit this power to about 1.0 mW, or less; the precise value is not important as long as this calibration power is low enough that essentially no change in junction temperature occurs. The voltage across A and B (V_{cbo}) can now be read. V_{cbo} will be approximately 0.5 V with the circuit shown and with the junction temperature at about 70°C. The 10-turn precision potentiometer is adjusted until the bucking voltage of the 1.354-V mercury cell matches V_{cbo}, and the oscilloscope reads zero volts with switch S3 thrown to the potentiometer wiper contact. Since the voltage across the potentiometer has been adjusted to 1.000 V (with a digital voltmeter), the potentiometer reads V_{cbo} to the nearest millivolt. This precision is mandatory, since the change in V_{cbo} is 2.0 mV/°C of junction temperature change.

b. Power loading of transistors under test. After each transistor has been calibrated, and the potentiometer settings recorded, the chassis (Fig. 5) is removed from the oven and allowed to cool to room temperature. With the potentiometer set at the calibrated reading, switch S1 is closed, and power is supplied to the transistor, thus heating the junction and raising its temperature above that of the chassis. The voltage across A and B is now approximately 1.5 V, and the power being delivered to the transistor is of the order of several hundred milliwatts. The switch S3 shorts out the oscilloscope during this loading, so that the vertical preamplifier, which has been set to read a vertical sensitivity of 1.0–5.0 mV/cm, will not be overdriven.

The junction temperature stabilizes several seconds after load power is delivered. When it has, the relay-operated switches S2 and S3 are thrown simultaneously. The junction immediately starts to cool; and, as it does, the oscilloscope traces the change in V_{cbo}. The power supplied to the transistor is adjusted until the cooling curve traced on the oscilloscope begins at zero, indicating that the junction before cooling was at the calibrated temperature. This procedure was used in order to observe the rate of change of V_{cbo} versus time. The slope of the V_{cbo} versus time curve can be used to estimate the degree of radiation cooling present, since, in effect, this is merely a calibrated cooling curve.

Figures 3 and 4 show cooling curves with the power adjusted so that the junction temperature is at the calibrated temperature. These curves quantitatively show the nature of the transient thermal characteristics of the transistor junction region. The initial cooling rate is greater than exponential, indicating that the effect of radiation cooling is appreciable, for if conduction cooling alone were present, the curve would be purely exponential.

c. Transistors tested. The following transistor types were tested:

Transistor type	Case type
2N1132	TO-5
2N1613	TO-5
2N1973	TO-5
2N995	TO-18
2N915	TO-18
2N2222	TO-18

The majority of test runs were made with the chassis at room temperature in free air. One test run was made

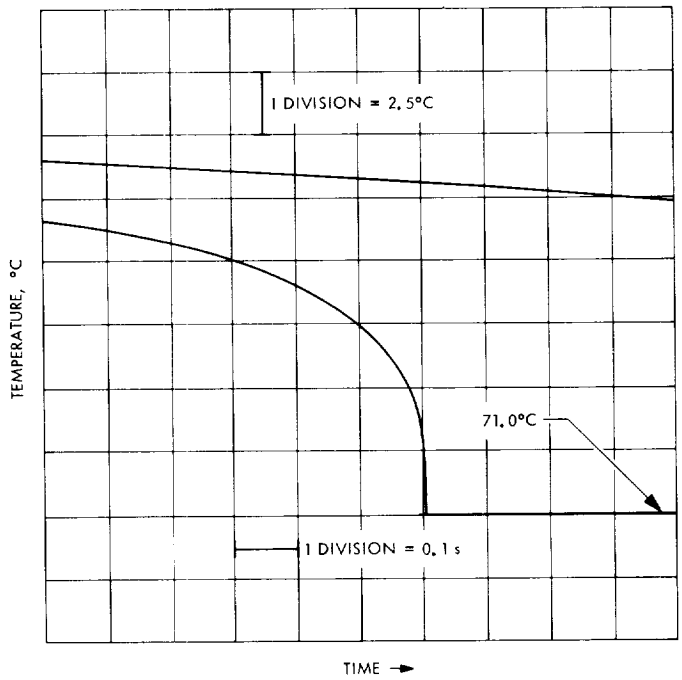


Fig. 3. Cooling curve for transistor 2N1613

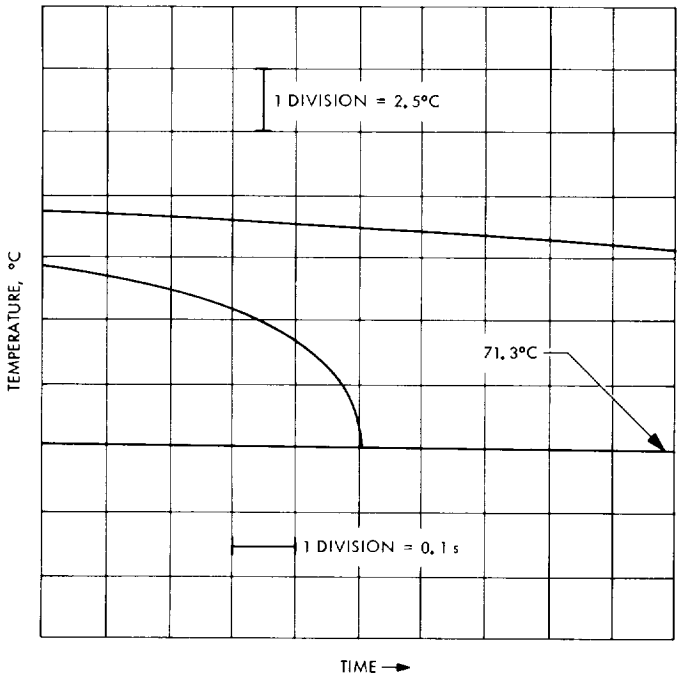


Fig. 4. Cooling curve for transistor 2N1973

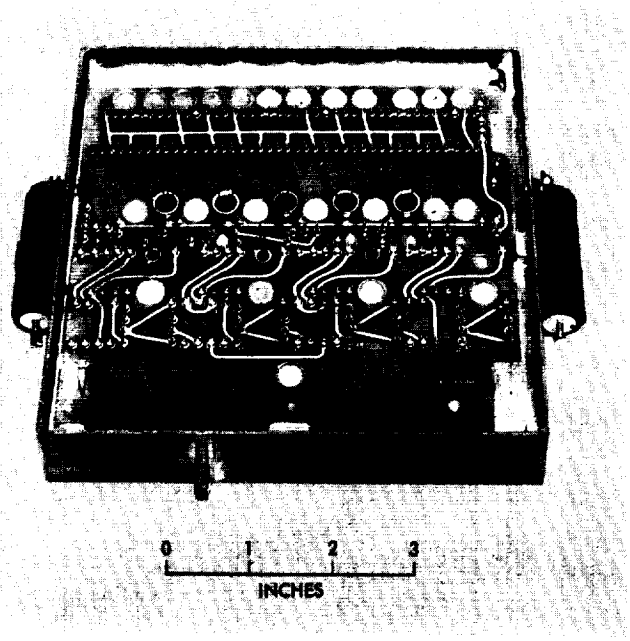


Fig. 5. Test chassis for transistor thermal resistance tests

in an 18-in. bell-jar vacuum chamber, which verified that the vacuum has little effect on the transistor dissipation. A 3 to 5% derating factor applied to free-air dissipation values is in substantial agreement with results from the vacuum chamber.

Since the thermal resistance is being measured between the junction and the chassis, it is implicitly assumed that the chassis is an isothermal heat sink; that is, that T_{ch} exists as a single measurable temperature. To justify this assumption, thermocouples were attached to the chassis beneath three transistors to check for hot spots. They indicated local temperatures between 1.0 and 2.0°C higher than ambient temperature. Since the mass of the entire chassis is large compared to the mass of the portion of the chassis affected by this local heating, the assumption of isothermal chassis temperature seems reasonable and conservative, since it results in a higher value for R_t than might actually be the case.

Table 3. Measured thermal resistance for several transistors

Transistor type	JEDEC case type	R_t, °F/mW		T_{ch}, °F	$T_j - T_{ch}$, °F
		Air	Vacuum		
2N1132	TO-5	0.145		73.4	86.5
			0.140	75.2	85.0
		0.137		73.4	100.2
			0.138	75.2	98.2
2N1132	TO-5	0.130		73.4	87.0
			0.143	75.2	85.0
		0.130		73.4	100.0
			0.149	75.2	96.5
2N1132	TO-5	0.112	—	72.5	86.3
		0.117	—	72.5	101.0
2N1132	TO-5	0.124	—	71.6	88.2
		0.123	—	71.6	101.7
2N1613	TO-5	0.144		73.4	85.5
			0.152	75.2	83.7
		0.140		73.4	101.7
			0.147	75.2	100.0
2N1613	TO-5	0.144		73.4	87.0
			0.152	75.2	85.3
		0.138		73.4	99.8
			0.142	75.2	96.2
2N1973	TO-5	0.143		73.4	87.0
			0.158	75.2	85.2
		0.133		73.4	99.8
			0.148	75.2	96.2
2N2222	TO-18	0.300	—	72.5	101.0
2N915	TO-18	0.295	—	72.5	101.5
2N995	TO-18	0.236	—	72.5	87.0

4. Summary

The measured values of thermal resistance for each transistor tested are given in Table 3. Most of the tests were run with two calibration points on the V_{cbo} versus temperature curve, since the cooling curves indicated some radiation heat transfer inside the transistor cans. Those devices in TO-18 cans showed very little difference between air and vacuum tests and are reported to be the same in both test conditions. The highest value of thermal resistance between junction and chassis measured for a transistor in a TO-5 case was approximately 0.15 °F/mW which, when translated into the amount of power dissipation per temperature rise, is approximately 6 mW/°F. The highest value of thermal resistance measured for transistors in a TO-18 case was approximately 0.30 °F/mW which, when translated into power dissipation/temperature rise, is 3 mW/°F.

Although these measured values are less by a factor of two than that predicted by the mathematical model, the mathematical model is conservative, as desired. The mathematical model could probably yield closer numbers if the physical constants were corrected.

The difference in thermal resistance between TO-5 and TO-18 cans appears to be predominately a function of the heat-transfer area from the lid of the can to the chassis.

Reference

1. Schneider, P. J., *Conduction Heat Transfer*, Addison-Wesley Pub. Co., Inc., Reading, Mass., 1955.

B. Simplifying Complex Miniature Interconnections, *L. Katzin*

1. Introduction

This article presents an approach to overcome some of the undesirable features associated with short-run complex miniature interconnections. This complexity decreases standardization and reliability and increases lead time and cost. A unique machine has been designed which permits small quantities of complex miniature assemblies to be intraconnected at efficiencies and a reliability level normally associated with mass production. Two packaging methods which achieve these efficiencies are possible with this machine.

2. Concentric Electrode Welding Machine

Complex miniature interconnecting involves a large number of simple operations: routing, joining, and terminating of an electrically conducting medium. These operations have been automated almost entirely with complex wire-wrap machines. This complexity was tolerable for short-run production but not for miniature interconnecting. The two changes that are necessary are process simplification and miniaturization.

The machine designed and built for complex miniature interconnecting performs in a semi-automated mode.

The need for this machine became obvious when short-run complex assemblies were first interconnected with multilayer boards. The number of modifications which were required, the high cost, the long lead time, and the questionable reliability, even before the modifications were incorporated, were all valid reasons for searching for alternate approaches. Wire wrap was too large, since the miniature assembly required terminating on 0.050-in. centers. A technique to which JPL had contributed in the early development stages—the magnet wire-welding technique—was considered. This technique uses a conventional opposed-electrode welder, in which the upper,

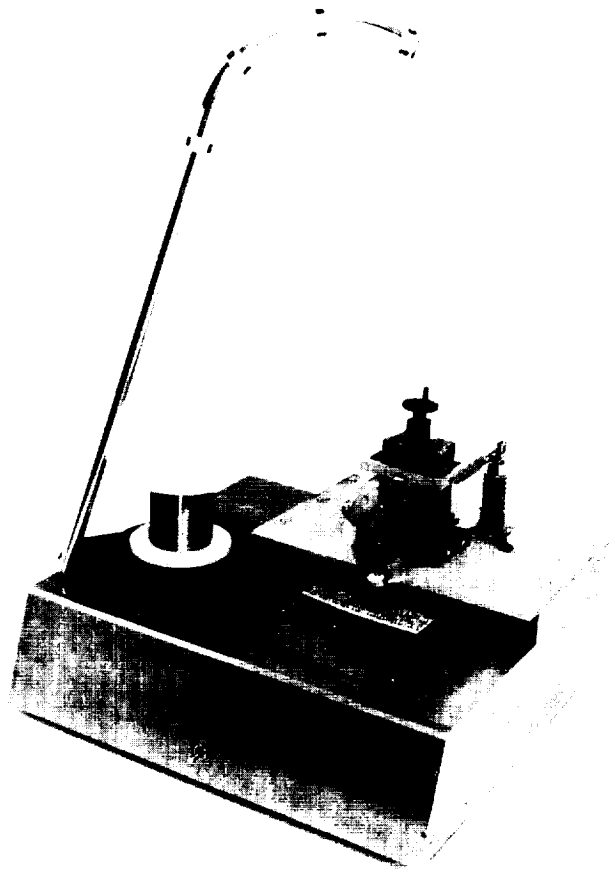


Fig. 6. Concentric electrode welding machine

or upper and lower, electrodes are heated to a temperature that will plasticize the insulation on the magnet wire and cause it to flow until physical contact is made between upper electrode, bare wire, and lower electrode. The weld is then effected as with an uninsulated wire. This approach had some advantages but also had the following shortcomings:

(1) Two electrodes, a terminal, and a wire had to be aligned on a single axis.

(2) The hot electrode could burn the insulation from adjacent wires.

(3) Because the insulation thickness on the wire varies, the amount of heat required to melt the insulation will also vary. This will change the wire temperature, which will change the electrical conductivity of the wire. These changes will require adjustments in weld schedules, depending on insulation thickness.

The new machine (Fig. 6) has removed the above disadvantages, and by so doing has made it possible to adapt this machine to fully automatic short-run capability.

The basic concept of the machine is shown in Fig. 7. Rather than using heat to locally displace the insulation, pressure is used. The amount of pressure required to break through the insulation is considerably greater than the welding pressure; therefore, two cycles are used. The first cycle applies the heavy pressure for piercing the insulation, and the second cycle applies a lesser pressure for welding. The two pressures are mechanically independent, and each has its own adjustment. Figure 8 shows a No. 34 AWG nickel wire (0.0063-in. diam) with Teflon insulation, welded to a gold-plated type 302 stainless terminal, 0.036-in. in diameter.

There is no bottom electrode in the machine, as the two electrodes are concentrically telescoped. The outer electrode makes a spring collet contact to the terminal shank, while the inner electrode (which is a hollow capillary) feeds the wire and presses it onto the top of the terminal. Since the electrode assembly feeds the wire, it also routes the wire. The routing is accomplished by moving the work from point to point. Unlike wire-wrap interconnection, a series of joints to terminals are accomplished with a single length of wire.

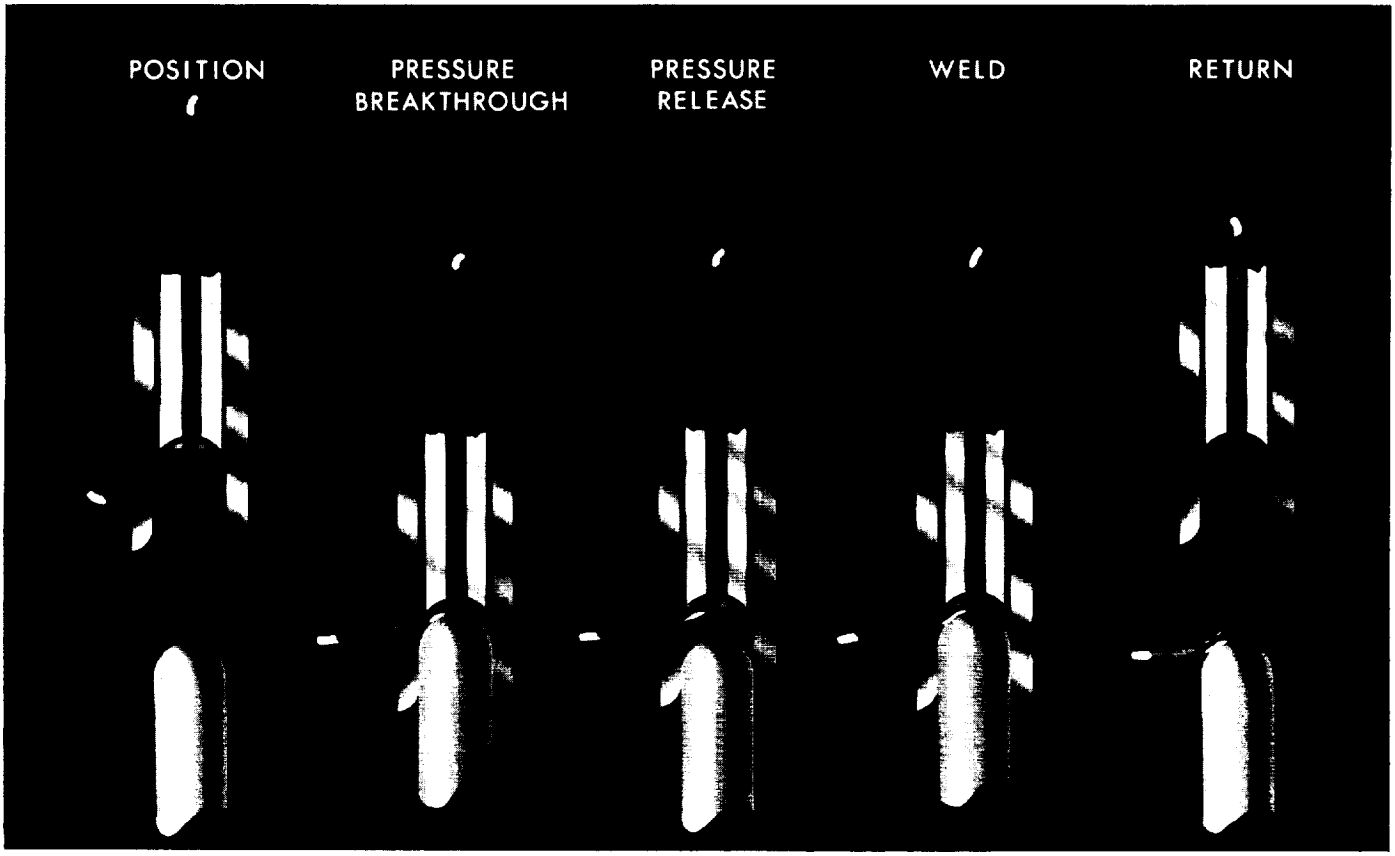


Fig. 7. Concept of concentric electrode welding

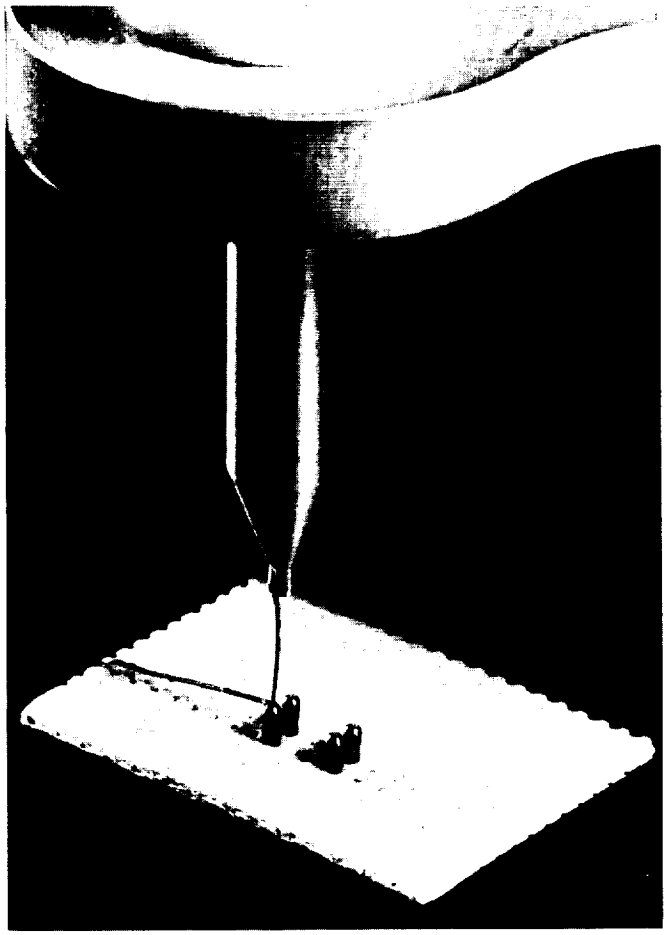


Fig. 8. Weld made on concentric electrode welder

The outer electrode serves a triple function. Its first function is to make a good electrical contact to the shank of the terminal. Its second function is to pilot the electrode assembly onto the terminal. This second function is very important, because, it not only makes the alignment problem less critical and more perfectly repetitive, but it also prevents the inner electrode from deflecting during the pressure cycles. The third function of the outer electrode is to clear the top of the terminal of all previously routed wires. This is implemented by the normal down stroke of the electrode. The top of the terminal is spherical. As the outer electrode descends, it pushes all adjacent wires out of the way prior to enveloping the straight portion of the terminal shank. This eliminates any possibility of a previously routed wire interfering with the implementation of the current weld. Because the top of the terminal is spherical (convex), the lip of the inner electrode is also spherical, but in a matching concave geometry.

The insulated wire emerges from the center of the inner electrode and trails in the direction of the previous weld. Before the electrode descends to make a weld, the wire is moved by a mechanical finger into the slit of the outer electrode. This prevents the wire from getting pinched between the outer electrode and the terminal shank. By moving the wire into the outer electrode slit, since the outer electrode is rotationally stable, all welds will be made with the wire approaching the terminal from the same direction. This has the advantages of providing an extra service loop in most point-to-point connections, providing better symmetry to the panel and also making automatic wire cutoff much easier. To date, the automatic cutoff is still in the drawing stage; however, once the machine is automated, which is the next step, this function will become mandatory.

The automation of the machine will be accomplished by the addition of a numerically controlled table to transport the work. The table has a speed of 100 in./min. A 6 × 6-in. panel with an average run length of 2 in. and a welding speed of 1 sec (the machine is driven by a 60-rpm motor, with one revolution per weld cycle), would produce over 1600 welds per hour, not counting wire cutoff time, which will probably be a 1-sec operation, based on an additional motor revolution for the cutting action. This would amount to interconnecting approximately one hundred 14-lead integrated circuits per hour. The interconnections are point-to-point; therefore, programing time is very short. There is no art work, no laminating, no alignment problem, and successive units made from the same punched tape will be identical. Repairs or modifications are quick, simple, and reliable, and such changes can be easily programmed into the punched tape or cards.

The carrier for the wiring and the components can be either planar or three-dimensional modular. Both methods have been developed, and each has its own advantages and unique applications.

3. Planar Carrier

The planar carrier consists of a perforated board, and gold-plated terminals. The cost of materials and labor required for assembly of the perforated boards is a small fraction of that for the multilayer boards.

The terminals that are used have two interesting features: (1) As shown in Fig. 9, the terminals have a self-locking feature which permits them to snap into the board and remain permanently in place, but yet removable, without the use of staking tools or adhesives. (2) The U-shaped terminal has a double leg which provides two wire attachment posts per component lead. Since wires can be joined

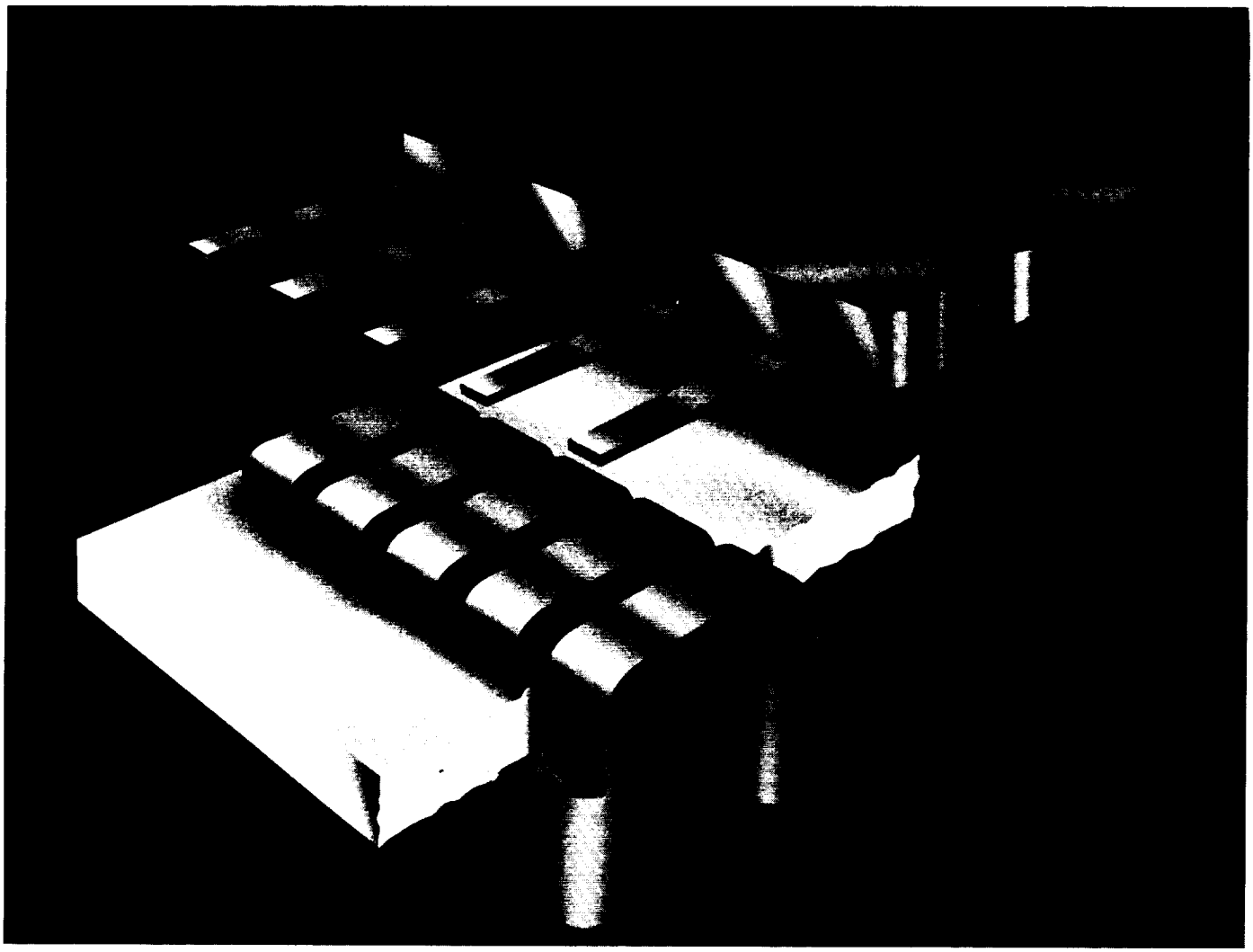


Fig. 9. U-terminal

to a terminal without being terminated, each post has the equivalent capacity of two wire-wrap connections, or a total of four equivalent wire-wrap terminations per component lead. The back of the terminal is the integrated-circuit attachment surface, and the leads may be parallel-gap-welded or reflow-soldered to the terminal.

4. Folded Stick Module

Another method of using the concentric electrode welder for high-density integrated-circuit packaging is the folded-stick module. The interconnecting method is the same as for the planar packaging; however, by using the folded stick geometry, three advantages are derived: (1) The wiring is protected within the fold of the stick. This forms a self-contained mold if the wiring needs to be encapsulated. (2) The component density on a system level is improved. The stick module has a component density of thirty 14-lead integrated circuits per cubic inch, and relatively little of this is lost in the subsystem packaging. (3) A very simple back panel can be used to combine many sticks in a system. This back panel need not be more than a one-sided printed circuit board.

Figure 10 shows how the folded stick module is wired in the flat and then folded to conceal the wires. The integrated circuits are attached to the two outside surfaces of the stick, and the output terminals extend from the stick along the folding line. The folded stick module is molded from polypropylene, a thermoplastic which permits the two molded-in hinges to be an integral part of the module.

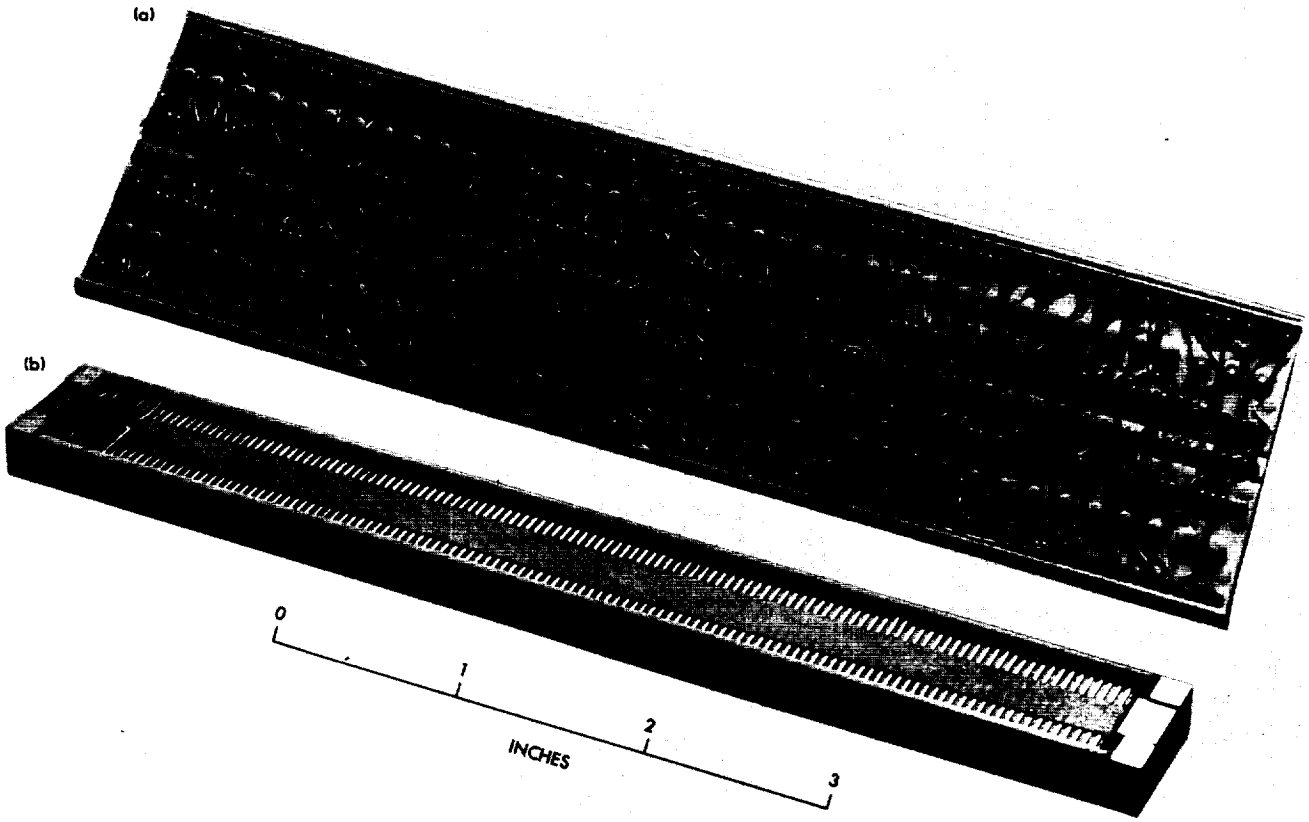


Fig. 10. Folded stick (a) wiring side (b) closed

5. Conclusions

The complex interconnecting problem has been reduced to a series of simple, high-dexterity, high-volume operations. Short run automation is thus made possible.

C. Documentation of Wiring Harnesses Using Punched-Card Techniques, *W. G. Kloezeman*

1. Introduction

Punched IBM cards can be used in conjunction with a card sorter and a card printer to document wire harnesses. This technique was developed in an effort to save time and to increase the flexibility of the present wiring documentation method which requires that a draftsman draw a pictorial representation of each connector. This drawing shows each pin of the connector, its wiring destination, the wire size, the function of the wire, and any pertinent twist or shield information. A wiring change requires that the drawing be corrected and then reissued, which is time consuming and costly. Also, the drawings fade over a period of time and the hand lettering can at times become illegible.

The basic idea behind this punched card technique is that each pin of each connector is represented by a punched card. A splice in this documentation method is treated as if it were a single pin connector. Thus, the number of punched cards for a harness equals the number of connector pins in the harness plus the number of wires in the harness that route to splices. All the cards representing a connector or splice are punched, sorted on a card sorting machine, and then used to print wiring documentation for that connector or splice. Also, wiring changes are easily accommodated. All cards involved in the wiring change are separated from the other cards; new cards are punched; and the wiring documentation is reprinted to obtain the latest configuration. There is no waiting for drawing vellums to be updated, and fabricating to engineering change orders or to obsolete prints is eliminated.

2. Procedure

The first step in using this technique is to punch the required input information onto the IBM cards in the 80 card columns. The information includes: the harness number, connector-pin or splice, wire destination, wire size, twist and shield requirements, wire length, wire type, signal or power direction, impedance, maximum signal voltage, sensitivity of the wire to noise, circuit data sheet number, wire function, general note reference, date of last card change, and harness revision letter.

The cards are then machine-sorted according to harness, sorted again according to connectors or splices within each harness, and the stacks arranged in consecutive order according to pin.

Next, the cards are ready to be run through an IBM 407 card printer to obtain wiring harness documentation. This consists of the information punched on the cards plus information coming from the character emitter of the printer. The printing format has 120 spaces available in comparison with the 80 columns on the punched cards, allowing room for the insertion of blanks.

A heading card which titles the information printing and a connector information card which describes the connector are inserted at the front of each connector stack before printing. After the cards are printed, the connector and splice sheets representing each harness are then stapled together. A wiring change is easily accommodated. The cards involved in the wiring change are removed, and new cards are punched. New wiring documentation for each affected connector is then printed and stapled with its respective harness. Figure 11 shows the documentation for a five-pin connector, and Fig. 12 shows the documentation for a splice.

9W109 2052A1P1 ST 10090 5S BAY 5 2.4 KHZ HEATER POWER 04/68 A

9WHRN	CONN	PIN	CONN	PIN	WS	TW	S	SH	R	WL	T	D	IMP	VTG	NI	CDS NO	FUNCTION	N	DT/68 A
9W109	52A1P1	1	54A1P1	1	22	15	2	10		23	X	R	121	155	113 3	004 014 C16	PWR 2.4 KHZ TO AG		04/68 A
9W109	52A1P1	2	59W1J3	B-	22					27	R	R	272	501	232 2	008 031 G05	G2 IRS COOLDOWN RTN		04/68 A
9W109	52A1P1	3	SPL 10		24					35	X	L	352	661	622 2	008 031 G05	G2 IRS COOLDOWN EXC		05/68 A
9W109	52A1P1	4	SHD RTN	A	24			10	A	2			15B				SHD 10 SHD RTN A		04/68 A
9W109	52A1P1	5	54A1P1	2	22	15	2	10		23	R	L	101	105	83 3	004 014 C16	PWR 2.4 KHZ TO AG RTN		04/68 A

HARNESS NUMBER · CONNECTOR · PIN · CONNECTOR · PIN · WIRE SIZE · TWIST REFERENCE NUMBER · TWISTED GROUP SIZE · SHIELD NUMBER · SHIELD RETURN LETTER · WIRE LENGTH · WIRE TYPE · SIGNAL OR POWER DIRECTION · IMPEDANCE · MAXIMUM SIGNAL VOLTAGE · NOISE INDEX NUMBER · CIRCUIT CLASS · CIRCUIT DATA SHEET NUMBER · FUNCTION · GENERAL NOTE REFERENCE · DATE · REVISION LETTER

Fig. 11. Documentation for a five-pin connector

9W109 SPLICE NO 1 BAY 3 PYRO HARNESS 03/68 A

9WHRN	SPL NO	CONN	PIN	WS	TW	S	SH	R	WL	T	D	IMP	VTG	NI	CDS NO	FUNCTION	N	DT/68 A
9W109	SPL 1	54A1P1	11	24			6		23	X	R	132	150	172 2	008 012 K07	T2 SOLAR PANEL DEPLOY		03/68 A
9W109	SPL 1	52A0P1	3	24			6		31	X	L	171	150	172 2	008 012 K07	T2 SOLAR PANEL DEPLOY		03/68 A
9W109	SPL 1	59W1J3	A-	24			6		14	X	L	111	150	172 2	008 012 K07	T2 SOLAR PANEL DEPLOY		03/68 A

HARNESS NUMBER · SPLICE NUMBER · CONNECTOR · PIN · WIRE SIZE · TWIST REFERENCE NUMBER · TWISTED GROUP SIZE · SHIELD NUMBER · SHIELD RETURN LETTER · WIRE LENGTH · WIRE TYPE · SIGNAL OR POWER DIRECTION · IMPEDANCE · MAXIMUM SIGNAL VOLTAGE · NOISE INDEX NUMBER · CIRCUIT CLASS · CIRCUIT DATA SHEET NUMBER · FUNCTION · GENERAL NOTE REFERENCE · DATE · REVISION LETTER

Fig. 12. Documentation for a splice

3. Conclusion

The use of this technique in documenting wiring harnesses has the following advantages:

(1) The quality of the documentation is improved. Wire harness drawings that fade are eliminated and hand lettering is not required.

(2) The flexibility of the documentation is increased. The cards can be used to print wiring harness documentation that is tailored to the requester's needs. For example, the cards could be sorted to obtain a stack of cards representing all twisted wire pairs in a particular harness by sorting for this characteristic in its respective card column.

(3) A wiring change is easily handled. The latest wiring information is immediately available.

IX. Solid Propellant Engineering

PROPULSION DIVISION

A. Study of Radiative Cooling of a Graphite Composite Nozzle, *S. Fogler*

1. Introduction

For the design study on radiatively cooled nozzles for a long-burning solid propellant rocket motor containing 760 lb of propellant, it became necessary to estimate the temperature profile and heat flux at the outer nozzle surface. This need arises from the potential heating problem that may occur in spacecraft components such as solar panels or hydrazine tanks which are adjacent to a midcourse propulsion system. If the nozzle surface temperatures prove to be so high that thermal insulation for the spacecraft components becomes impractical or the required insulation weight becomes prohibitively high, the incentive for developing a light-weight radiatively cooled nozzle will no longer exist.

The radiatively cooled nozzle under study is fabricated from a composite consisting of a filament-wound graphite fiber in a graphite matrix, i.e., Carbitex. In the calculations, it was assumed the nozzle would have a throat diameter of 2.405 in. and an expansion ratio of 69, it would be radiating solely to deep space, the initial nozzle temperature would be $-260°$F before motor firing, and the nozzle emissivity was taken to be 0.9. The stagnation temperature T_0 of the combustion gases was estimated at 4600°F, and the maximum motor burn time was 200 s. Estimates were made on the temperature profile in the solid nozzle as well as on the heat flux from the nozzle.

2. Theory

Under the conditions stated above, the differential equation describing the variation of temperature T with distance down the nozzle x and time t for a nozzle thickness which is much less than the nozzle radius is given by:

$$\rho C_p l \frac{\partial T}{\partial t} = h(T_g - T) + lk \frac{\partial^2 T}{\partial x^2} + \frac{kl}{r} \frac{\partial T}{\partial x} \frac{\partial r}{\partial x} - \epsilon \sigma T^4$$

where

T = temperature, °F

t = time, h

r = radius of nozzle at any distance x, ft

σ = Stefan–Boltzmann constant, BTU/ft²-h-°R⁴

k = thermal conductivity of nozzle, BTU/ft-h-°R

l = nozzle thickness, ft

C_p = heat capacity of nozzle, BTU/lb-°R

ρ = density of nozzle, lb/ft³

ϵ = emissivity

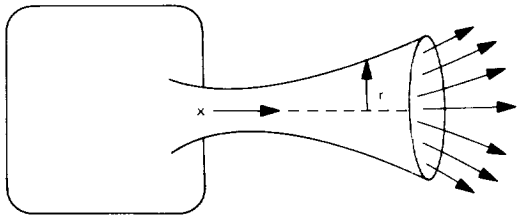


Fig. 1. Composite nozzle schematic

A schematic diagram of the nozzle is shown in Fig. 1. Utilization of the equation for the heat transfer coefficient h, developed by Bartz (Ref. 1), will give approximate values for the propellant under consideration, but within the accuracy required:

$$h = 559\left(\frac{2.4}{6.36 + 2x\tan\alpha}\right)^{1.8}\sigma_1$$

where α = the angle the nozzle cone makes with the center line, and σ_1 = the dimensionless factor accounting for variation of gas density and viscosity across the boundary layer which is a prescribed function of x, $0.6 \leqq \sigma_1 \leqq 1.6$. The nozzle thickness is on the order of 0.08 in. As a result of an order of magnitude analysis, it can be shown that axial conduction in the nozzle can be neglected, and that the thermal capacity is quite small. Consequently, a steady-state nozzle wall temperature is reached in a small fraction of the total burning time.

3. Computational Technique and Results

At each point x down the nozzle, the surface temperature was determined by an iteration procedure on an IBM 7094 computer. In the iterative technique used, convergence was set so that if the calculated temperature was within 5 deg of the predicted temperature, the iteration loop was terminated after convergence was achieved. The distance downstream from the nozzle mounting was incremented and a new heat transfer coefficient was calculated. With this new value of h, the iterative procedure was repeated and the temperature at this new x calculated. By continuing in this manner, one can determine the temperature T as a function of distance x downstream from the nozzle mounting. This relationship between T and x is shown in Fig. 2. Also shown on this plot is the heat flux q, which

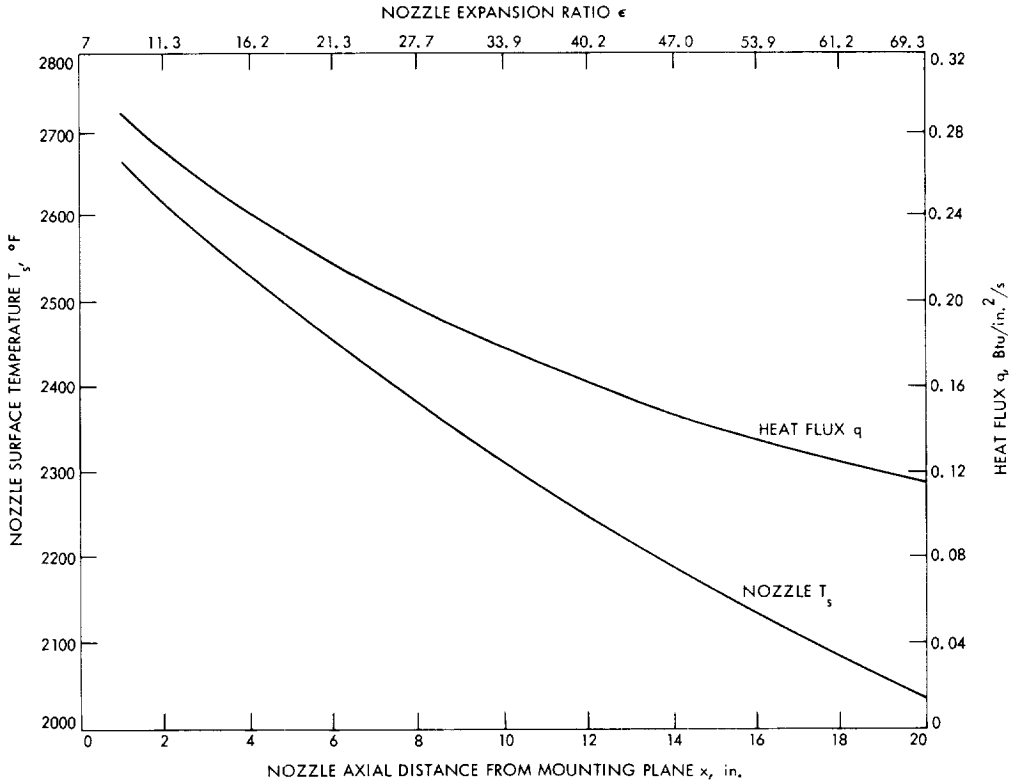


Fig. 2. Temperature vs distance downstream from nozzle mounting

is readily calculated from the equation $q = h\ (T_g - T)$ once T is determined. The temperature of the nozzle varies from approximately 2700°F at the nozzle mounting to 2040°F at the exit of the nozzle. The average temperature of the nozzle is quite close to 2350°F, and the average heat flux is on the order of 0.2 BTU/in.^{2}/s.

In another segment of this study, a calculation was performed to determine whether or not the solar panel would be made inoperative by the effects of radiative energy. If it is assumed that only 10% of the radiative energy leaving the nozzle hits the solar panels, a quick engineering estimate shows the solar panels will exceed their maximum allowable operating temperature in approximately 20 s.

As a result of this work, refined calculations will be made of the energy radiated to all spacecraft components and methods will be examined for reducing the nozzle-radiated heat flux to these components. This reduction in heat flux could be accomplished by submersion and isolation of the nozzle by the rocket motor, or through the introduction of a thermal insulation such as a graphite felt shroud around the nozzle cone.

Reference

1. Bartz, D. R., *Jet Propulsion Laboratory External Publication 351*. Jet Propulsion Laboratory, Pasadena, Calif., July 31, 1956.

B. Solid Propellant Rocket Motor Command Termination by Water Injection, *L. D. Strand*

1. Introduction

Command termination of solid propellant motors by water quench has been demonstrated to be feasible in many tests using motors with propellant weights up to 19,000 lb. In these tests, some parameters have been varied in an attempt to define extinguishment criteria; however, all attempts to correlate the data have given questionable results. This is primarily due to, and points out, the inadequate understanding of the quench mechanism. In these correlation attempts, it has been assumed that: (1) rapid cooling of gases causes a dp/dt sufficient for extinction, (2) cooling of the gases below some threshold temperature lowers heat feedback to the propellant below that necessary for self-supported combustion, and (3) a water film covers the entire surface of the propellant, thereby quenching combustion. Because the extinguishment mechanism is not adequately understood, it is difficult to make confident estimates of necessary water quantities, optimum injection rate, and optimum injection techniques to accomplish termination. A complete understanding of the means necessary to prevent reignition of the propellant is also lacking.

2. Test Program

A study of water quench of solid rockets has been initiated to: (1) determine the optimum method of water injection to obtain complete extinguishment, (2) better understand the quench mechanism so that it may be optimized, (3) establish means for predicting water requirements of any given motor, and (4) determine feasibility and performance of water extinguishment of flight motors.

The program consists of two phases: (1) a laboratory motor phase, which is currently underway, and (2) a small solid-propellant motor phase. The laboratory motor, which is fully described in *Subsection 3*, consists of a slab burning window motor which is capable of accepting several different types of water injectors. In this phase, a knowledge of the efficiency of each of the different types of injector and spray patterns and an indication of the dominant quench mechanism, or mechanisms, will be sought. Using this information, an attempt will be made to develop a theoretical model capable of correlating the data of each injector and providing general extinguishment criteria.

The objective of the Phase 2 portion of the program is to gain confidence in the Phase 1 extinguishment criteria and the reliability of this type of impulse control. It will consist of the combustion termination by water injection of a 5-in.-diameter by 6-in.-long cylindrical grain motor with a propellant weight of approximately 3½ lb and a flight-weight cylindrical motor with a low length–diameter ratio and a propellant weight of 60 lb. Knowledge of the optimum values of the quench parameters gained in the slab burning motor will be applied in these tests to determine any motor configuration and size dependency.

3. Phase I Test Apparatus

A block diagram of the original version of the Phase 1 water quench apparatus is shown in Fig. 3. The principal components, shown on the horizontal line of the diagram, consist of a high-pressure nitrogen source, a water accumulator, a flow meter, an explosively operated valve, and a window motor. The side leg is used to

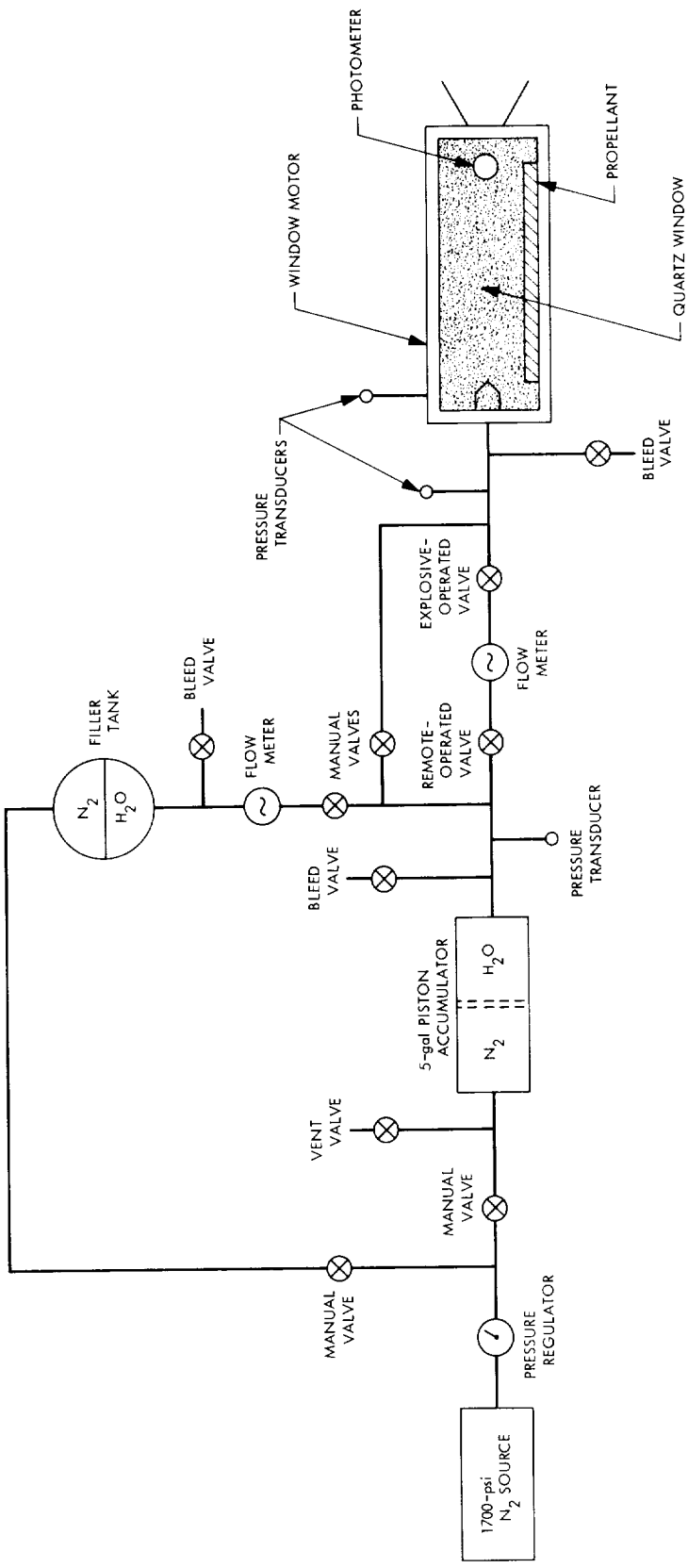


Fig. 3. Experimental water quench apparatus

purge all air out of the line and to fill the accumulator with the amount of water to be injected. Valves are provided to enable all air to be bled out of the system. A photograph of the system is shown in Fig. 4. Several modifications to the system were found necessary when attempts were first made to terminate the motor. These will be discussed in *Subsection 7*.

A schematic of the laboratory motor is shown in Fig. 5. It is fabricated of 304 stainless steel. The overall dimensions are: length, 17.2 in.; width, 3.75 in.; and height, 4.5 in. Flat slabs of propellant, 13.7 in. long, 1.75 in. wide, and 0.50 in. thick, cast on steel plates, are mounted on either the bottom surface only or both the top and bottom of the motor. The slabs are restricted on all four sides

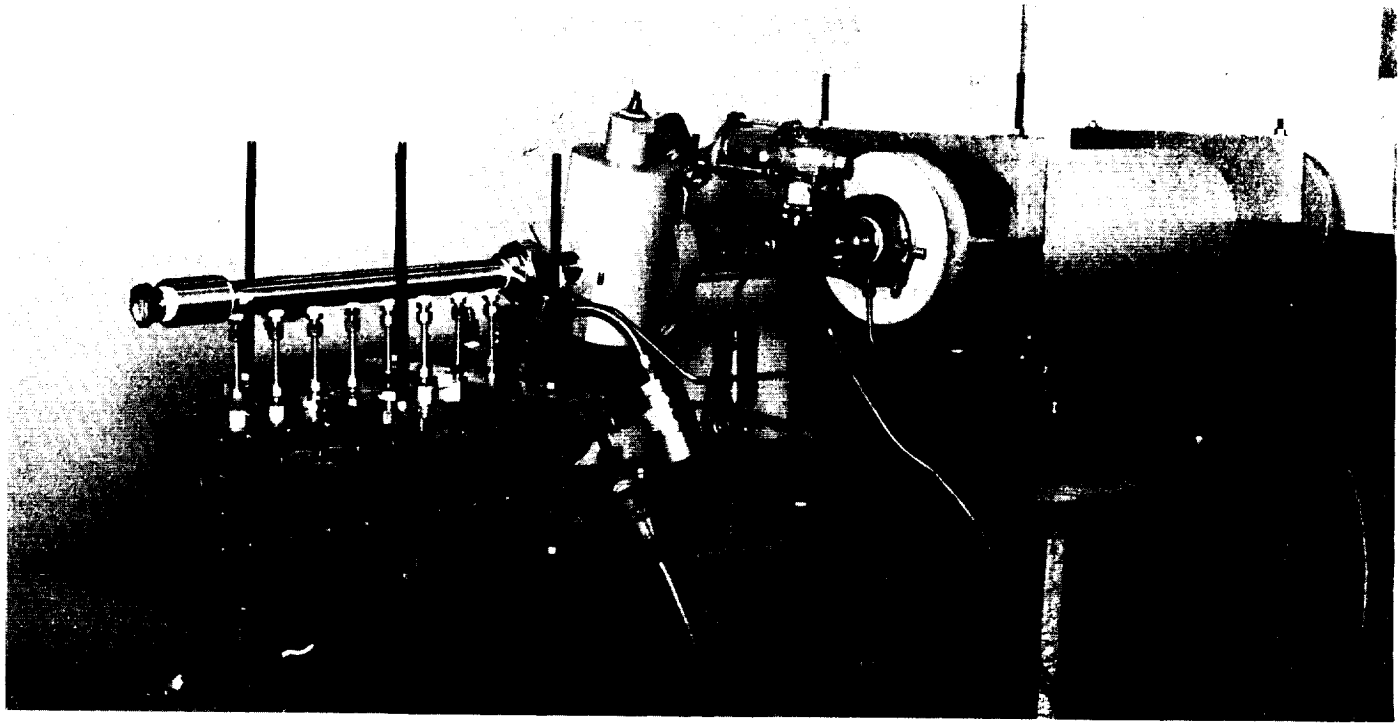


Fig. 4. Water quench test system

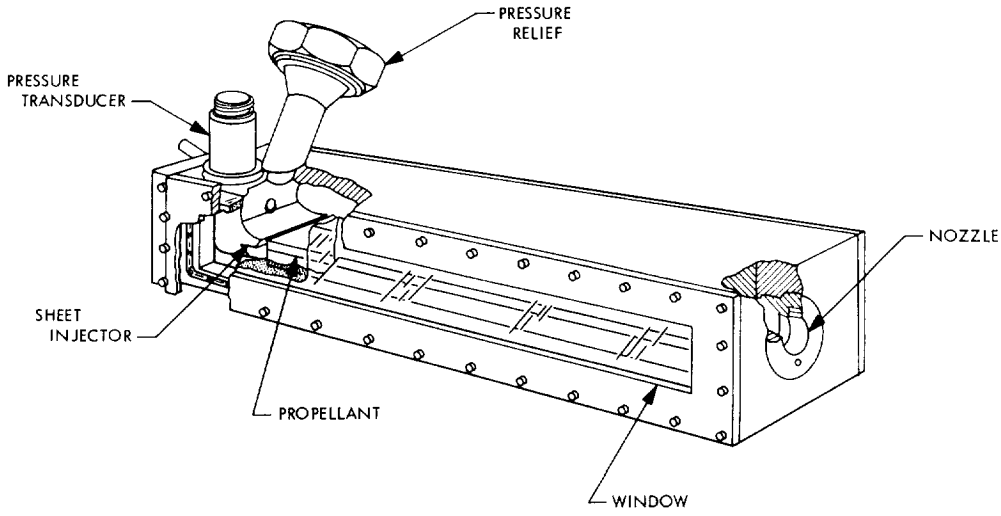


Fig. 5. Laboratory window motor

and have a burning area of approximately 24 in.^{2}. The threaded nozzle is solid copper and converges to its exit throat diameter. The exit diameter can be varied from 0.400 to 1.000 in. The motor was designed to use $\frac{5}{8}$-in.-thick quartz windows, although plexiglas windows have been used for a majority of the tests. The windows seat on copper gaskets and are sealed with rubber O-rings. As a guard against over-pressurizing the motor, the motor has a 300-psig pressure relief diaphragm.

The motor is designed to accept three different types of water injector systems. Figure 4 shows the first basic system, a set of eight $\frac{1}{4}$-in. overhead injectors. Only the lower propellant slab is used for this system. Solid-cone, hollow-cone, flat jet, and atomization-type injectors will be used. The second system, a head-end sheet injector, is shown in Fig. 5. This injector design, based on the work reported in Ref. 1, is designed to lay a thin sheet of water over the lower propellant slab burning surface. The third system consists of a $\frac{3}{4}$-in. head-end injector. Solid-cone, hollow-cone, and atomization-type injectors will again be used. A different head-end closure is used for each of the igniter systems: a simple flat plate for the over-head system and plates for mounting the head-end sheet and $\frac{3}{4}$-in. injectors. The motor accepts a Dynisco water-cooled and a Taber pressure measurement transducer. Both are of the strain gage type. Motor ignition is provided by a hot-wire igniter system, with the igniter leads passing through the motor nozzle (Fig. 4).

The explosive valve is manufactured by Cartridge Actuated Devices, Inc. It is of the normally closed type. Opening time is a few milliseconds, and is accomplished by gases generated from an electrically initiated pyrotechnic squib, forcing a cutter bar to shear off an aluminum closure.

A 1-in. Waught turbine-type flow meter, model FL-1658, measures the water flow rate. A solenoid-operated globe valve is included in the water line for checkout and injector spray tests, where millisecond opening times are not required. The water accumulator, manufactured by Greer Hydraulics, Inc., is a 5-gal floating-piston type. As shown in Figs. 3 and 4, Taber pressure gages measure the water pressure at the accumulator exit and upstream of the injector(s).

4. Instrumentation

Instrumentation for each test consists of the accumulator and injector water-pressure Taber gates, the motor-pressure Taber and Dynisco gages, the turbine flowmeter, a silicon diode photovoltaic cell to indicate cessation of combustion, and high-speed motion-picture photography. The Dynisco pressure transducer, model PT49C-5C (0–500 psig), has a flat ($\pm 10\%$) frequency response to 10,000 Hz (Ref. 2). The photocell is mounted on the motor window in approximately the position shown in Fig. 3.

A 16-mm Hycam camera is used for the motion-picture coverage. The camera has a full-frame maximum speed of 11,000 frames/s and a half-frame rate of 22,000 frames/s.

The pressure transducers, flowmeter, and photocell outputs are amplified and recorded on a 7-channel Ampex FR 600 tape recorder. Signals indicating the closing of the hot-wire igniter, camera power, and explosive valve squib circuits are also recorded on one of the channels. The tape recorder is operated at a recording speed of 60 in./s and played back on a Consolidated Electrodynamics Corp. oscillograph at a play-back speed of 6.5 in./s, resulting in an 8/1 expansion of the time scale.

5. Test Procedure

The standard test procedure is as follows: The window motor and explosive valve are assembled, using the motor exhaust nozzle required to provide the desired chamber pressure, and installed in the test system (Fig. 4). During assembly of the motor, all water injectors are plugged with Celvacene vacuum grease. The pressure gages are calibrated and installed. With all bleed valves open, the accumulator is slightly pressurized with nitrogen to drive the piston to its fully expelled forward position. The filler tank is filled with water and water is flowed under nitrogen pressure to the portions of the main water line forward and aft of the explosive valve until the line is purged of all air. The main line bleed valves are closed and the filler tank is filled with the amount of water to be injected into the motor. The nitrogen side of the accumulator is vented to atmosphere and the water in the filler tank is flowed under pressure into the accumulator. The accumulator nitrogen vent valve is closed, the electrical leads to the hot-wire igniter, camera, and explosive valve are connected, and the accumulator is pressurized with nitrogen to the desired pressure level.

Now the test can be performed. The tape recorder is started and the hot-wire igniter circuit is closed. When the propellant is visually observed to be completely ignited, the camera power circuit is closed. The closing of the camera circuit starts a digital sequence timer that closes the explosive valve circuit after a time delay required for the camera to accelerate up to its nominal operating speed (approximately $1\frac{1}{2}$ ms). Following opening of the explosive valve, nitrogen pressure forces the

floating piston to the forward end of the accumulator, injecting the predetermined quantity of water into the motor. The rate of injection is determined by the area of the injectors and the accumulator nitrogen pressure.

6. Motor Checkout and Development

In order to (1) verify that the over-head injector system was producing a uniform spray pattern along the motor and (2) determine the spray pattern for each type of injector nozzle planned to be used in the test program, a series of cold-flow motor injector water spray tests were performed. The spray patterns in the motor were photographed using both still and motion-picture photography. The data verified that the spray pattern was uniform and that obtaining complete coverage of the propellant surface would be no problem with any of the injectors.

The first motor firings had the primary purpose of (1) checking the combustion stability characteristics of the motor and the quartz windows and (2) developing a propellant igniter system and a propellant side-restricter. A restricter was desired that would inhibit the sides of the slab from burning, but one that would regress with the burning top surface so as not to interfere with the camera view of the burning surface. Ten motors have been fired to date, with water injection employed in the last five.

As a precautionary measure, steel plates were used in place of the windows in the first two tests. No erratic pressure oscillations were experienced throughout the two firings. No instability problems have been encountered in any of the motor firings conducted to date.

The quartz windows were used in the third test. Fine cracks were found near the edge of one of the windows upon post-fire inspection. The window eventually burst during the fifth test firing. For all subsequent tests, the quartz windows were replaced with double-layer plexiglas. A ⅛-in. inner sheet was replaced after each test. Ablation of the inner windows does obscure detail viewing, but the plexiglas should be adequate for a majority of the tests. Quartz windows may again be used for a few of the final tests, when specific details are desired.

A simple hot-wire igniter was used in the first five tests. This resulted in excessively long ignition delays of several seconds. An igniter paste technique was tried in the next two tests. It consisted of a pyrotechnic igniter paste that was spread over the top surface of the slab and allowed to harden. The paste was ignited with a hot wire. The paste proved to be difficult to ignite, and the technique gave greater ignition delays than were obtained using the hot-wire igniter. In the eighth test, an igniter system consisting of three hot-wire propellant igniters wired in parallel was used. A relatively rapid ignition for a hot-wire system was obtained (½-s delay). The same igniter was used in the ninth test, but the use of too large an igniter propellant weight resulted in over-pressurization of the motor and bursting of the motor pressure-relief valve. The igniter system was modified to two hot-wire igniters wired in parallel for the tenth test, providing satisfactory ignition of the motor.

The restricter tried in the first two tests was cellulose dissolved in methyl ethyl ketone (MEK) solvent. This is the standard restricter used at JPL for restricting the strands of propellant used in burning-rate bomb tests. In both tests, the restricter failed less than a second after the motor reached its operating pressure, as indicated by a sudden rise in the chamber pressure. In the third test, a Sil-Guard cement was tried as a restricter. Post-fire inspection showed that the Sil-Guard had not regressed, but remained intact although badly charred. For the fourth test, MEK thickened with more cellulose was used and was marginally successful. A material called Okun's Original Liquid Plastic Vinyl was tried as a restricter in the fifth test, and it seemed to perform as desired. Armstrong cement was tried in the sixth test. It did not prove successful; there was some charring, but no regression. The Okun vinyl material was used as the propellant restricter for all subsequent tests.

7. Initial Extinguishment Tests

Five water quench tests have been performed (tests 6–10), with successful termination in three cases. Details of the tests are given in Table 1. The results of tests 8 and 10 will be discussed in greater detail later in this subsection.

Motion-picture data were obtained in all the tests. The camera field of view included the entire length of the propellant slab in the sixth test. This was reduced to approximately the forward half of the slab for the subsequent tests. Ektachrome EF-B and EF-Daylight color film was used. Different lighting setups and camera framing rates were tried. For the first three water quench tests, the lighting consisted of six 500-W RSP2 lights, four for front lighting and two as back lights. This proved inadequate to view the combustion process during termination, when self-illumination from the combustion falls off rapidly. For the remaining tests, the lighting system consisted of three 750-W reflector spot back lights focused

Table 1. Test parameters and results

Test	Nozzle throat diameter, in.	Injector system	Accumulator water volume, ml	Accumulator water pressure, psig	Nominal motor pressure, psig	Total burning time, s	Camera average frame rate, frames/s	Test results
6	0.660	8 flat spray	120	300	75	3.2	4000	No extinguishment
7	0.660	8 flat spray	150	90	80	0.05	4000	Extinguishment
8	0.600	8 flat spray	90	300	95	2.6	8000 (half frame)	Extinguishment
9	0.600	8 hollow cone	60	1400	—	0.2	8000	Motor Over-pressurized
10	0.640	8 hollow cone	140	2000	70	2.0	8000	Extinguishment

along the front half of the propellant slab. The film was over-exposed in the focus region of the spots, but more detail was obtained during the quench process. This will be discussed further under the description of the tenth test.

The overhead water injector system (Fig. 4) was used in all the tests. In the first three tests, the injectors used were of the flat spray type, with an orifice diameter of 0.070 in. and a nominal spread angle of 40 deg. The nozzles were aligned so that the jets spread along the centerline of the propellant slab. Hollow-cone atomizing nozzles were used for the last two tests. Nozzle diameter was 0.028 in.

Injection problems were encountered in these initial tests. No flow meter signal was obtained in the sixth test, and examination of the motion-picture film showed that only a minute amount of water had been injected. Apparently, when the explosive valve plug was sheared off, most of the injectant water flowed into some unfilled void in the line instead of being injected into the motor. The obvious culprit was the valve itself, which has some void space inside it. To correct this, the valve was rotated 180 deg from the position shown in Fig. 4, and the top of the valve was drilled and tapped and a bleed valve installed. This allowed the valve to be purged of air when the portion of the water line downstream of the valve was purged.

In the seventh test, the accumulator was pressurized to 300 psig, but the water pressure dropped to approximately 90 psig before the test could be initiated. The decision was made to go ahead and run the test with this low water-injection pressure. The motor extinguished in spite of the low water pressure and resultant low rate of injection. To find the reason for the water pressure drop, the water side of the accumulator was disassembled. A significant void was found in the transition region from the accumulator exit to the 1-in. water line, a portion of which was not being completely bled of air prior to test. When the accumulator was pressurized with nitrogen, the small amount of water in the accumulator was forced into this void, and the piston bottomed against the end of the accumulator. A bleed valve was installed in the water exit end of the accumulator, only partially correcting the problem.

For the eighth test, the camera was driven at the same voltage, but the half-frame prism was used, doubling the framing rate. Prior to test, the accumulator was filled with 90 ml of water and pressurized to 300 psig. Before the test could be initiated, the water pressure dropped to 30 psig. This time the lines were repurged and the accumulator refilled and again pressurized to 300 psig. The water pressure stayed at this value until all water was injected. The test resulted in extinguishment of combustion. The extinguished propellant slab is shown in Fig. 6. The roughness of the surface is due to the three hot-wire igniter propellant pieces that were stapled to the slab surface. Ignition of the propellant under the igniter pieces was apparently delayed. No photocell data were obtained, but the motion-picture film indicated that extinguishment appeared to be very rapid, occurring within 6–8 ms after the onset of water injection. The maximum rate of depressurization calculated from the oscillograph record was approximately 6,000 psi/s. The flow meter response seemed to be surprisingly good, indicating flow a few milliseconds after the explosive valve signal and giving an integrated total water volume injected that agreed quite well (within 10%) with the volume loaded into the accumulator. The average flow rate during quench, as measured by the flow meter, was 4.8 lb/s. The total amount of water required to terminate

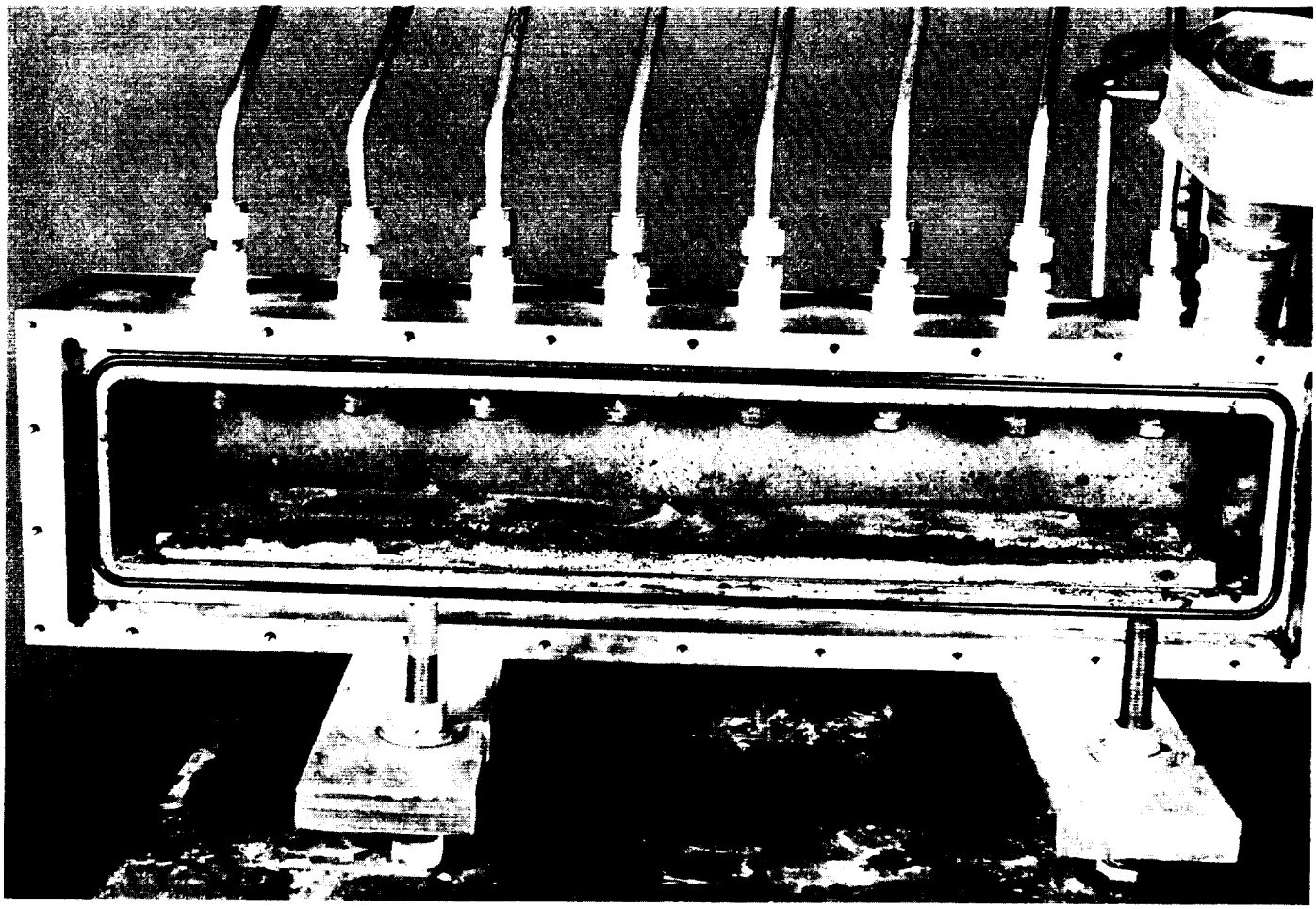


Fig 6. Extinguished propellant slab

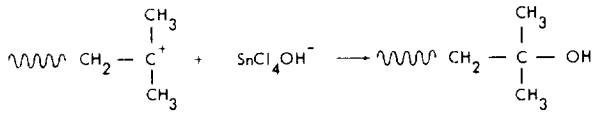
combustion, again as measured by the flow meter, was 0.03 to 0.05 lb (14 to 23 ml).

In tests 9 and 10, the full-frame prism was again used, but the camera drive voltage was increased to give an estimated average framing speed of 8,000 frames/s. The water injector nozzles were changed to hollow-cone atomizing nozzles. For the ninth test, the accumulator was filled with approximately 60 ml of water and pressurized to 1700-plus psig. Because of gas voids, the accumulator piston again bottomed, and the water pressure dropped significantly before the test could be initiated. As previously described, the motor overpressurized at ignition in this test, bursting the motor pressure-relief diaphragm before the water quench system could be activated.

To avoid bottoming of the piston in test 10, the accumulator water volume was raised to approximately 140 ml of water. The water was pressurized to 2000 psig, and the pressure held constant throughout the test. The test resulted in termination of combustion. Figure 7 shows the extinguishment portion of the oscillograph record. The oscillations in the injector water pressure are probably due to the nature of the injector manifold. The calculated maximum depressurization rate was approximately 7500 psi/s. Extinguishment appeared to begin approximately 4 ms after the onset of water injection and was complete after 10 ms, as indicated by both the photocell and motion-picture data. The measured water flow rate during quench was 14 lb/s. The total amount of water injected up to the onset of extinguishment was 0.06 lb (28 ml).

The framing rate in the region of quench was 6000 frames/s. As mentioned earlier, the intense back lighting provided more than adequate illumination. The motion-picture record showed the following: Combustion appeared to continue normally for approximately 4 ms following appearance of the water spray. At 4 ms the

double bonds per chain. This tends to indicate that the diradical mechanism takes place when VO Cl_3–naphthalene catalyst is used. Unfortunately, termination by some reactions (other than proton abstraction) apparently occur to produce non-olefinic endgroups irrespective of the predominant initiation or propagation reactions. One of these reactions is known to be of the type shown below (Refs. 3, 4).



Thus, rigorous elimination of water and use of an appropriate cocatalyst initiator may result in the formation of difunctionally unsaturated polymers.

One method of modifying poly(isobutylene) is to react it with bromine or chlorine and then dehydrohalogenate to produce double bonds. Although chlorination has not been investigated, there is reason to believe that substitutive bromination may occur predominantly near the chain ends. An isobutylene polymer with number-average molecular weight (Mn) of about 1000 was reacted with bromine and dehydrobrominated. It was then treated with ozone and reduced to convert the unsaturation to hydroxyl groups. The Mn of the final polymer was essentially unchanged and contained a high concentration of OH and no unsaturation.

After the dehydrobromination, the unsaturation, measured by infrared, corresponded to a functionality of approximately 2.7, which was roughly equivalent to the unsaturation initially present plus that expected from complete dehydrobromination of the excess bromine added. The final polymer also contained some carbonyl, apparently because of incomplete reduction after ozonolysis. Due to the small sample of polymer obtained, a reliable hydroxyl functionality determination has not been carried out. However, this polymer and a sample of hydroxyl-terminated saturated polybutadiene (SPBU) were reacted with pyromellitic dianhydride in xylene and the intrinsic viscosities were measured before and after chain extension for each polymer. The increase was relatively small but comparable in both cases: 58% for the hydroxylated poly(isobutylene) ester and 55% for the SPBU esterification product. This result may mean that the functionalities are approximately the same; however, a more definitive and reliable determination is required. Larger amounts of unsaturated poly(isobutylene) are being prepared both for further functionality analysis and to investigate other reactions for the modification of unsaturated groups, such as epoxidation.

References

1. Yamada, N., Shinada, K., and Hayashi, T., *J. Polym. Sci.*, Series B, Vol. 4, pp. 477–480, 1966.
2. McGuchan, R., and McNeill, I. C., *J. Polym. Sci.*, Series A-1, Vol. 4, pp. 2051–2062, 1966.
3. Dainton, F. S., and Sutherland, G.B.B.M., *J. Polym. Sci.*, Vol. 4, p. 37, 1949.
4. Biddulph, R. H., Plesch, P. H., and Rutherford, P. P., *J. Chem. Soc.*, p. 275, 1965.

B. Investigation on Sterilizable Polymer Battery Separators, Part II, *E. F. Cuddihy and J. Moacanin*

1. Introduction

Sterilizable battery separator materials have been prepared from a graft copolymer of polyethylene and poly(potassium acrylate) (Ref. 1). An investigation of their properties is being carried out in order to develop a set of physical and chemical criteria to characterize both the starting polyethylene and the separator materials (SPS 37-50, Vol. III, pp. 166–169). This article provides additional information on the structure and properties of these materials.

2. Molecular Weights

The number-average and weight-average molecular weights of the starting polyethylene (PE) film (Petrothane 301) were found by gel permeation chromatography to be 15,000 and 48,200, respectively, giving a M_w/M_n ratio of 3.18; this ratio is a measure of the spread in molecular weight.

3. Additives

An attempt was made to detect and identify additives in the polyethylene by exposing the materials to a vacuum and trapping the volatiles in a liquid nitrogen condenser. In addition to water and a trace of CO_2, the presence of an esterified material was detected by infrared analysis.

4. Radiation Exposure

The graft copolymer battery separator is prepared by first crosslinking the PE with divinylbenzene and then grafting with acrylic acid. Gamma radiation from a Cobalt 60 source is employed for both reaction steps, and

the PE is exposed to a cumulative dosage of around 2½ mrad. The PE sheets were exposed to this level of radiation both in air and vacuum. After irradiation, both materials remained completely soluble in α-chloro naphthalene, indicating the absence of crosslinking. These results lead to the conclusion that the crosslinking in the final battery separator material is definitely due to the divinylbenzene, and possibly, to some extent, to the grafted acrylate chains.

The crystalline content of both irradiated samples was unchanged, whereas the crystalline content of the PE in the battery separator was found to be decreased by 7% (SPS 37-50, Vol. III). Previously, it was suggested that the decrease in crystallinity could have been caused by radiation exposure, or from partial melting caused by heating the separator to 80°C during an extraction step to remove excess acrylic acid homopolymer. Eliminating radiation exposure as a possibility implies that the crystalline content was reduced by the partial melting. However, since recrystallization did not occur when the separator was cooled back to room temperature, some permanent change had to occur during the extraction step. Perhaps some post-grafting occurred in the newly created amorphous region produced by the partial melting. In any event, if it is desirable to reduce crystallinity in the battery separator, perhaps the simple expedient of an increase in the extraction temperature to effect more extensive melting would be in order.

Finally, the only effect noted for the PE at this level of radiation occurred for the air-exposed sample. Infrared analysis revealed the presence of carbonyl groupings not present in the starting materials, showing that some oxygen uptake occurred during irradiation.

5. Divinyl Benzene

Divinyl benzene (DVB) is employed as a crosslinker in the battery separator. Exhaustive washing in hot benzene showed no free DVB present in the final product. Infrared analysis showed DVB present with either both or one of the vinyl groups reacted. Estimates of the concentration of these two, from infrared data, run about 2 wt % for the totally reacted DVB and 0.2 wt % for the partially reacted DVB.

6. Potassium Concentration Gradient

The battery separator material (designated GX-119) presently under study was prepared by immersing PE into a grafting solution containing acrylic acid monomer, and then exposing the system to gamma-radiation from a Cobalt 60 source. A total of 700 ft of PE (rolled into a cylinder) was employed for the grafting reaction. After irradiation, the film was washed in hot alcoholic KOH to extract out acrylic acid homopolymer and to convert the acid form to chains of poly(potassium acrylate) (PKA). Analysis of samples of the final product taken from the outside, middle, and inside of the roll revealed PKA concentration of 27, 43, and 38%, respectively. The concentration gradient of PKA along the radius of the roll of PE may be the result of the balance between the decreasing level of ionizing radiation with depth and the diffusion of fresh acrylic acid monomer (which is uniform throughout the roll) into the PE.

A high level of radiation impinging on the outside may generate free radicals faster than they can be consumed by the diffusing acrylic acid monomer. This would promote undesirable side reactions, e.g., disproportionation leading to the formation of acrylic acid homopolymer. As the radiation is attenuated by passing through the roll, its level may become sufficiently reduced to result in free radical formation occurring at a rate commensurate with the diffusion of acrylic acid monomer and, concomitantly, favor the grafting reaction.

7. Water Absorption

The graft copolymer is extremely hygroscopic and readily equilibrates with atmospheric moisture. A plot of equilibrium water absorption versus relative humidity for the GX-119 material is given in Fig. 1. It is seen that these materials pick up 13 to 14 wt % of water at a relative humidity, typical for this locale, of 40 to 50%.

The previous article (SPS 37-50, Vol. III) gave differential scanning calorimetry (DSC) patterns for the battery separator and the starting PE. Both PE and the separator exhibited an endotherm corresponding to the crystalline PE melting at around 110°C, while the separator exhibited a second unexpected endotherm maximizing at around 150°C. Infrared and ultraviolet analysis of the separator sample, before and after heating to these temperatures, revealed that the only change in the system was a loss of water. Analysis of the volatiles obtained under vacuum when heating samples to these temperatures showed the products to be predominantly water. Also, the second endotherm was absent in the DSC patterns taken on samples after the stripping of water. However, when samples were allowed to re-equilibriate with moisture, the second endotherm reappeared. Therefore, this

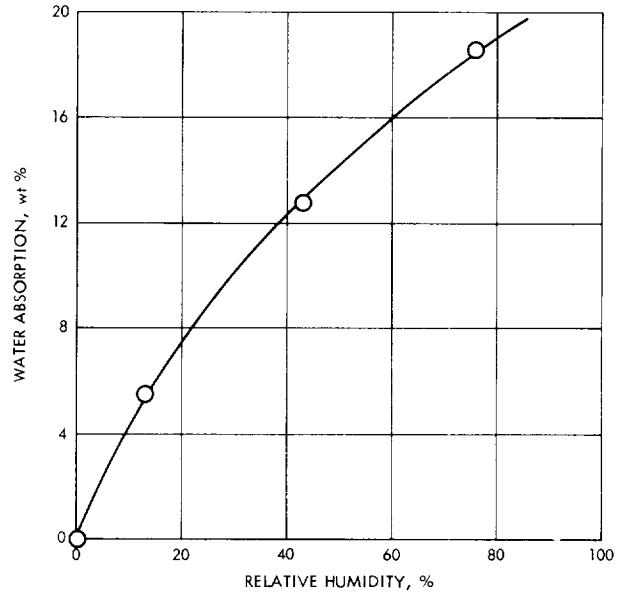


Fig. 1. Equilibrium water content of GX-119 battery separators

endotherm appears to represent the heat of hydration and was calculated at 2 kcals/mole water.

8. RAI-116 Battery Separator

A sample of a battery separator material, prepared by RAI Research Corporation, chemically similar to the GX-119 materials was investigated. Properties of this material, designated RAI-116, are summarized in Table 1. The PKA content of 32% is within the range observed for the GX-119 materials. The 34% crystallinity content for RAI-116 is significantly lower than the 43% value for GX-119. The lower solubility of RAI-116 (14% versus 21% for GX-119) in good PE solvents reflects the lower crystalline PE content.

Table 1. Grafted–crosslinked polyethylene film RAI-116

Property	Determination
Density at 22°C,[a] g/cm^3	1.218
Melting point,[b] °C	105
Poly (Potassium Acrylate), wt %	32[c]
Polyethylene, wt %	68
Poly (Acrylic Acid), wt %	Undetected
PE crystallinity,[d] %	34
Solubility in good acrylic solvents,[e] %	0
Solubility in good PE solvents,[f] %	14

[a]Determined from buoyancy in silicone oil.
[b]Determined by differential scanning calorimetry (DSC).
[c]Determined from K determination by flame photometry.
[d]Determined from DSC using Petrothane 301 as reference.
[e]Methanol and KOH–Methanol solutions.
[f]α-Chloro-Naphthalene.

9. Role of Crystallinity

The results obtained prior to this study show that the final battery separators have retained a considerable proportion of the initial crystallinity present in the starting PE. This means that all of the crosslinking and grafting has been confined to the amorphous part of the PE initially present. This necessarily leads to an inhomogeneous distribution of the grafted acrylate on the PE as well as some restriction to the amounts of the grafted material.

The battery separator membrane functions to keep separate the constituents of the respective half-cells of the battery, but must permit the ready permeation of the hydroxyl group (the charge carrying species common to both half-cells). As crystallinity in membranes retards permeation, the amount of material transported per unit time for a given membrane decreases noticeably as the crystallinity content increases.

Finally, the dimensional stability of the membrane can be seriously affected since large changes in dimensions are encountered when the very dense crystalline regions melt. Environment, temperature, and mechanical strain all influence the kinetics of recrystallization, and, hence, the rate at which the membrane returns to its former state after heating. If the membrane is first properly fitted for its application and then subsequently heated (as during sterilization), the accompanying dimensional changes may present a serious operational problem, especially if recrystallization proceeds slowly, or not at all, and results in permanent dimensional changes. Evidence for this latter point comes from DSC measurements on heat-sterilized specimens whose crystallinity after one week was about 70% of the initial crystallinity. The role of crystallinity in membrane performance should not be overlooked. The fact that the RAI-116 membrane with its lower crystallinity content improved the performance of test cell operation suggests that crystallinity should be kept at a minimal level.

Reference

1. Adams, L. M., Harlowe, W. W., Jr., and Lawrason, G. C., *Fabrication and Testing of Battery Separator Material from Modified Polyethylene*, Project 01-1842 Final Report, Southwest Research Institute, San Antonio, Texas, June 7, 1966.

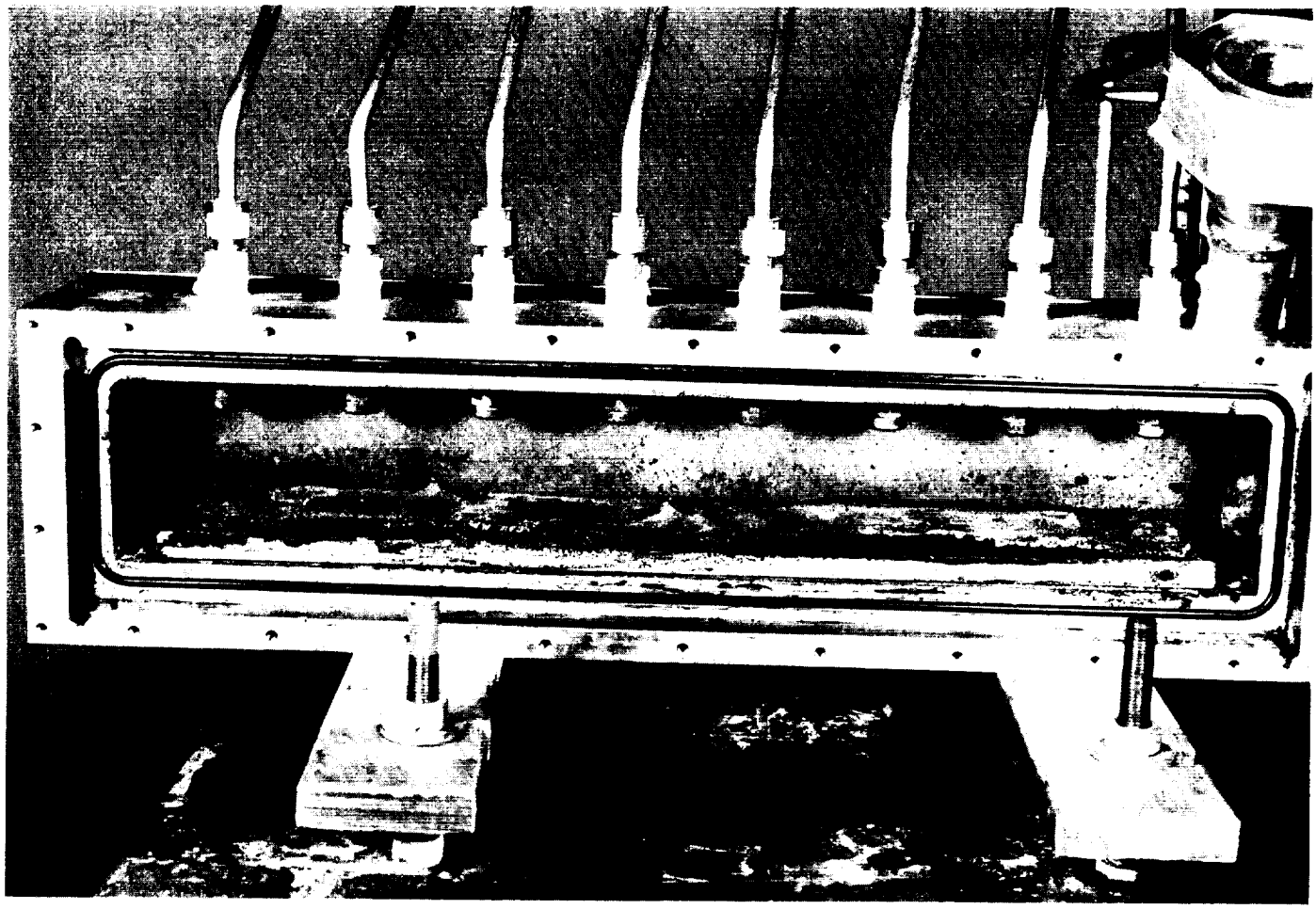


Fig 6. Extinguished propellant slab

combustion, again as measured by the flow meter, was 0.03 to 0.05 lb (14 to 23 ml).

In tests 9 and 10, the full-frame prism was again used, but the camera drive voltage was increased to give an estimated average framing speed of 8,000 frames/s. The water injector nozzles were changed to hollow-cone atomizing nozzles. For the ninth test, the accumulator was filled with approximately 60 ml of water and pressurized to 1700-plus psig. Because of gas voids, the accumulator piston again bottomed, and the water pressure dropped significantly before the test could be initiated. As previously described, the motor overpressurized at ignition in this test, bursting the motor pressure-relief diaphragm before the water quench system could be activated.

To avoid bottoming of the piston in test 10, the accumulator water volume was raised to approximately 140 ml of water. The water was pressurized to 2000 psig, and the pressure held constant throughout the test. The test resulted in termination of combustion. Figure 7 shows the extinguishment portion of the oscillograph record. The oscillations in the injector water pressure are probably due to the nature of the injector manifold. The calculated maximum depressurization rate was approximately 7500 psi/s. Extinguishment appeared to begin approximately 4 ms after the onset of water injection and was complete after 10 ms, as indicated by both the photocell and motion-picture data. The measured water flow rate during quench was 14 lb/s. The total amount of water injected up to the onset of extinguishment was 0.06 lb (28 ml).

The framing rate in the region of quench was 6000 frames/s. As mentioned earlier, the intense back lighting provided more than adequate illumination. The motion-picture record showed the following: Combustion appeared to continue normally for approximately 4 ms following appearance of the water spray. At 4 ms the

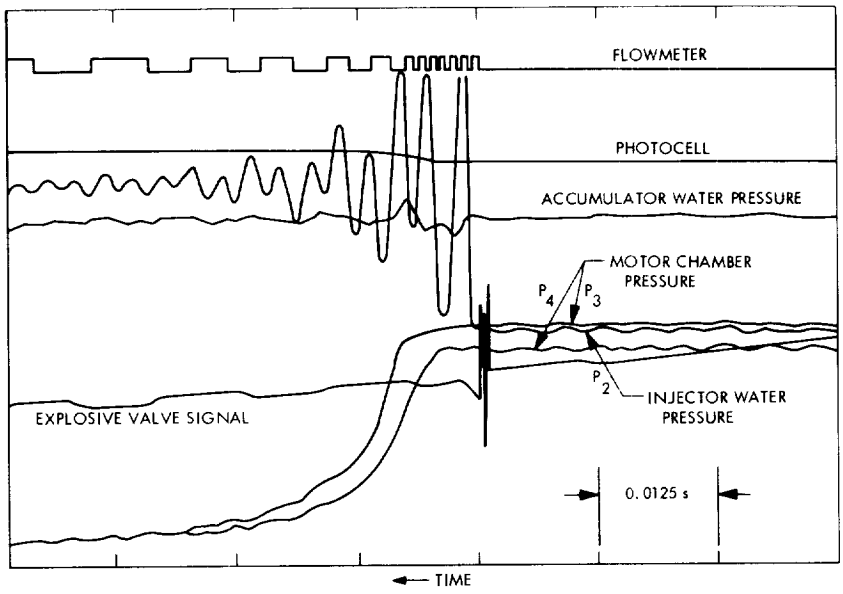


Fig. 7. Oscillograph record of test 10

combustion gases appeared to fly away from the propellant surface, as if the chamber had suddenly been vented. Light intensity in the motor then dropped off rapidly.

8. Discussion of Initial Results

Any conclusions drawn from the results of the few tests performed to date are, of course, very speculative. Nonetheless, certain distinctions between the different test results are apparent. Using the solid-cone water injector system, with its resultant direct impingement on the propellant surface (test 8), extinguishment was very rapid. It appeared to occur prior to any significant drop in the motor chamber pressure. The extinguishment had the characteristics expected of a simple thermal quench of the burning propellant. With the atomizing injectors (test 10), the extinguishment process was somewhat slower and coincided with the rapid drop in chamber pressure, resembling a rapid depressurization or dp/dt quench. Extrapolating the results of recent low-pressure L^* extinguishment studies (Ref. 3), the measured maximum rates of depressurizations for these tests are two to three times greater than that required to produce extinguishment at these low motor pressures. It is, therefore, hypothesized that, under low pressure conditions and for a propellant readily terminated by rapid depressurization, such a phenomenon may be the critical process in water quench. Further tests will have to be performed to substantiate this argument.

9. Future Effort

The piston accumulator used in the tests to date was sized for the Phase 2 full-scale motor test portion of the program. As has been described, the small quantities of water required to quench the laboratory motor do not displace the accumulator piston sufficiently to get a positive expulsion action when the accumulator is pressurized. A smaller (30 $in.^3$) accumulator has been procured and is being installed for use in the remainder of the Phase 1 effort.

All test techniques have been essentially defined. The Phase I quench tests will be continued, with the test variables consisting principally of water injector type and injection rate.

References

1. Riebling, R. W., *The Formation and Properties of Liquid Sheets Suitable for Use in Rocket Engine Injectors*, Technical Report 32-1112. Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1967.

2. Thomas, J. P., *Summary Technical Report on Transient Pressure Measuring Methods Research*, Contract NAS -36 and NAS 8-11216, Aeronautical Engineering Report 595 p. Princeton University, Princeton, N. J., Nov. 16, 1965.

3. Strand, L. D., *Summary of a Study of the Low-Pressure Combustion of Solid Propellants*, Technical Report 32-1242. Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1968.

X. Polymer Research

PROPULSION DIVISION

A. Saturated Hydrocarbon Prepolymers,

J. D. Ingham

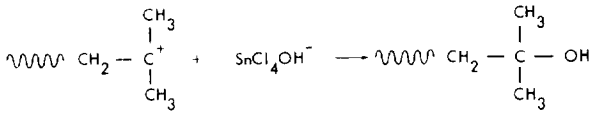
1. Introduction

This article summarizes the current status of a continuing effort (SPS 37-45, Vol. IV, pp. 113–115) on the synthesis of new saturated hydrocarbon prepolymers for use, primarily, as thermally and chemically stable propellant binders. The present effort is concerned with poly(isobutylenes). Although the initial work indicated that polymers containing terminal double bonds could be prepared directly by cationic polymerization in the presence of a suitable Lewis acid cocatalyst system, it has been established that none of the polymers prepared to date (in which adventitious water was the cocatalyst) contain more than one terminal double bond per chain. Two approaches are underway to attempt to increase the functionality to two. The first involves the study of other cocatalysts (under rigorously anhydrous conditions) that may lead to initiation and termination reactions producing terminal unsaturation to the extent that the functionality will be close to two. These studies will be discussed in a future SPS, Vol. III, article. The second approach, which is the primary subject of this article, involves chemical modification of prepolymers to obtain the desired endgroups.

2. Results and Discussion

For the polymerization of isobutylene by vanadium oxytrichloride in the presence of naphthalene, Yamada, Shimada, and Hayashi (Ref. 1) postulated a mechanism whereby an initiating cation would propagate at both ends. The cationic chain ends would then undergo termination by proton abstraction to give poly(isobutylene) with unsaturated chain ends. Although earlier analytical data indicated that such a polymer contained 1.6 unsaturated endgroups per molecule, more recent results indicate a value no greater than 1.1. The lower value is based on the amount of hydrogen consumed in a catalytic hydrogenation method and an infrared analysis calibrated with a polymer containing a known concentration of unsaturation. Both of these methods agreed with each other within less than 10%. Although the mechanism mentioned above may be partially operative, it does not occur exclusively and is not sufficient for the production of difunctional prepolymers. Analysis of a large number of $SnCl_4$–water (Ref. 2) and BF_3–water catalyzed poly(isobutylenes) showed that they contained 0.5 to 0.7 terminal

double bonds per chain. This tends to indicate that the diradical mechanism takes place when VO Cl_3–naphthalene catalyst is used. Unfortunately, termination by some reactions (other than proton abstraction) apparently occur to produce non-olefinic endgroups irrespective of the predominant initiation or propagation reactions. One of these reactions is known to be of the type shown below (Refs. 3, 4).



Thus, rigorous elimination of water and use of an appropriate cocatalyst initiator may result in the formation of difunctionally unsaturated polymers.

One method of modifying poly(isobutylene) is to react it with bromine or chlorine and then dehydrohalogenate to produce double bonds. Although chlorination has not been investigated, there is reason to believe that substitutive bromination may occur predominantly near the chain ends. An isobutylene polymer with number-average molecular weight (Mn) of about 1000 was reacted with bromine and dehydrobrominated. It was then treated with ozone and reduced to convert the unsaturation to hydroxyl groups. The Mn of the final polymer was essentially unchanged and contained a high concentration of OH and no unsaturation.

After the dehydrobromination, the unsaturation, measured by infrared, corresponded to a functionality of approximately 2.7, which was roughly equivalent to the unsaturation initially present plus that expected from complete dehydrobromination of the excess bromine added. The final polymer also contained some carbonyl, apparently because of incomplete reduction after ozonolysis. Due to the small sample of polymer obtained, a reliable hydroxyl functionality determination has not been carried out. However, this polymer and a sample of hydroxyl-terminated saturated polybutadiene (SPBU) were reacted with pyromellitic dianhydride in xylene and the intrinsic viscosities were measured before and after chain extension for each polymer. The increase was relatively small but comparable in both cases: 58% for the hydroxylated poly(isobutylene) ester and 55% for the SPBU esterification product. This result may mean that the functionalities are approximately the same; however, a more definitive and reliable determination is required. Larger amounts of unsaturated poly(isobutylene) are being prepared both for further functionality analysis and to investigate other reactions for the modification of unsaturated groups, such as epoxidation.

References

1. Yamada, N., Shimada, K., and Hayashi, T., *J. Polym. Sci.*, Series B, Vol. 4, pp. 477–480, 1966.
2. McGuchan, R., and McNeill, I. C., *J. Polym. Sci.*, Series A-1, Vol. 4, pp. 2051–2062, 1966.
3. Dainton, F. S., and Sutherland, G.B.B.M., *J. Polym. Sci.*, Vol. 4, p. 37, 1949.
4. Biddulph, R. H., Plesch, P. H., and Rutherford, P. P., *J. Chem. Soc.*, p. 275, 1965.

B. Investigation on Sterilizable Polymer Battery Separators, Part II, *E. F. Cuddihy and J. Moacanin*

1. Introduction

Sterilizable battery separator materials have been prepared from a graft copolymer of polyethylene and poly(potassium acrylate) (Ref. 1). An investigation of their properties is being carried out in order to develop a set of physical and chemical criteria to characterize both the starting polyethylene and the separator materials (SPS 37-50, Vol. III, pp. 166–169). This article provides additional information on the structure and properties of these materials.

2. Molecular Weights

The number-average and weight-average molecular weights of the starting polyethylene (PE) film (Petrothane 301) were found by gel permeation chromatography to be 15,000 and 48,200, respectively, giving a M_w/M_n ratio of 3.18; this ratio is a measure of the spread in molecular weight.

3. Additives

An attempt was made to detect and identify additives in the polyethylene by exposing the materials to a vacuum and trapping the volatiles in a liquid nitrogen condenser. In addition to water and a trace of CO_2, the presence of an esterified material was detected by infrared analysis.

4. Radiation Exposure

The graft copolymer battery separator is prepared by first crosslinking the PE with divinylbenzene and then grafting with acrylic acid. Gamma radiation from a Cobalt 60 source is employed for both reaction steps, and

the PE is exposed to a cumulative dosage of around 2½ mrad. The PE sheets were exposed to this level of radiation both in air and vacuum. After irradiation, both materials remained completely soluble in α-chloro naphthalene, indicating the absence of crosslinking. These results lead to the conclusion that the crosslinking in the final battery separator material is definitely due to the divinylbenzene, and possibly, to some extent, to the grafted acrylate chains.

The crystalline content of both irradiated samples was unchanged, whereas the crystalline content of the PE in the battery separator was found to be decreased by 7% (SPS 37-50, Vol. III). Previously, it was suggested that the decrease in crystallinity could have been caused by radiation exposure, or from partial melting caused by heating the separator to 80°C during an extraction step to remove excess acrylic acid homopolymer. Eliminating radiation exposure as a possibility implies that the crystalline content was reduced by the partial melting. However, since recrystallization did not occur when the separator was cooled back to room temperature, some permanent change had to occur during the extraction step. Perhaps some post-grafting occurred in the newly created amorphous region produced by the partial melting. In any event, if it is desirable to reduce crystallinity in the battery separator, perhaps the simple expedient of an increase in the extraction temperature to effect more extensive melting would be in order.

Finally, the only effect noted for the PE at this level of radiation occurred for the air-exposed sample. Infrared analysis revealed the presence of carbonyl groupings not present in the starting materials, showing that some oxygen uptake occurred during irradiation.

5. Divinyl Benzene

Divinyl benzene (DVB) is employed as a crosslinker in the battery separator. Exhaustive washing in hot benzene showed no free DVB present in the final product. Infrared analysis showed DVB present with either both or one of the vinyl groups reacted. Estimates of the concentration of these two, from infrared data, run about 2 wt % for the totally reacted DVB and 0.2 wt % for the partially reacted DVB.

6. Potassium Concentration Gradient

The battery separator material (designated GX-119) presently under study was prepared by immersing PE into a grafting solution containing acrylic acid monomer, and then exposing the system to gamma-radiation from a Cobalt 60 source. A total of 700 ft of PE (rolled into a cylinder) was employed for the grafting reaction. After irradiation, the film was washed in hot alcoholic KOH to extract out acrylic acid homopolymer and to convert the acid form to chains of poly(potassium acrylate) (PKA). Analysis of samples of the final product taken from the outside, middle, and inside of the roll revealed PKA concentration of 27, 43, and 38%, respectively. The concentration gradient of PKA along the radius of the roll of PE may be the result of the balance between the decreasing level of ionizing radiation with depth and the diffusion of fresh acrylic acid monomer (which is uniform throughout the roll) into the PE.

A high level of radiation impinging on the outside may generate free radicals faster than they can be consumed by the diffusing acrylic acid monomer. This would promote undesirable side reactions, e.g., disproportionation leading to the formation of acrylic acid homopolymer. As the radiation is attenuated by passing through the roll, its level may become sufficiently reduced to result in free radical formation occurring at a rate commensurate with the diffusion of acrylic acid monomer and, concomitantly, favor the grafting reaction.

7. Water Absorption

The graft copolymer is extremely hygroscopic and readily equilibrates with atmospheric moisture. A plot of equilibrium water absorption versus relative humidity for the GX-119 material is given in Fig. 1. It is seen that these materials pick up 13 to 14 wt % of water at a relative humidity, typical for this locale, of 40 to 50%.

The previous article (SPS 37-50, Vol. III) gave differential scanning calorimetry (DSC) patterns for the battery separator and the starting PE. Both PE and the separator exhibited an endotherm corresponding to the crystalline PE melting at around 110°C, while the separator exhibited a second unexpected endotherm maximizing at around 150°C. Infrared and ultraviolet analysis of the separator sample, before and after heating to these temperatures, revealed that the only change in the system was a loss of water. Analysis of the volatiles obtained under vacuum when heating samples to these temperatures showed the products to be predominantly water. Also, the second endotherm was absent in the DSC patterns taken on samples after the stripping of water. However, when samples were allowed to re-equilibriate with moisture, the second endotherm reappeared. Therefore, this

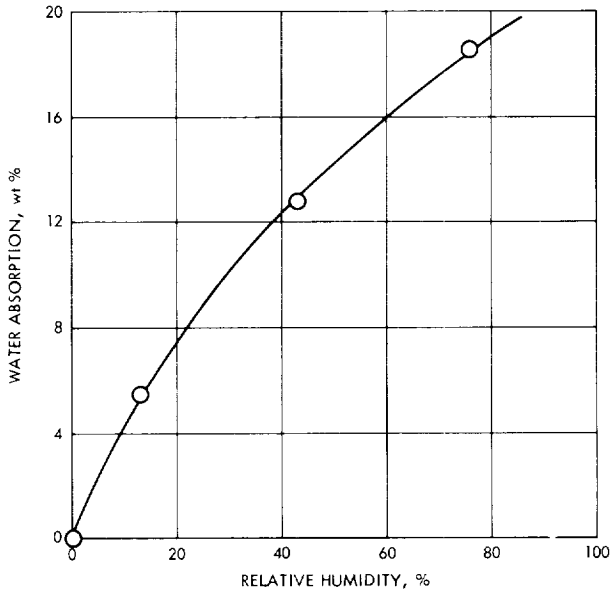


Fig. 1. Equilibrium water content of GX-119 battery separators

endotherm appears to represent the heat of hydration and was calculated at 2 kcals/mole water.

8. RAI-116 Battery Separator

A sample of a battery separator material, prepared by RAI Research Corporation, chemically similar to the GX-119 materials was investigated. Properties of this material, designated RAI-116, are summarized in Table 1. The PKA content of 32% is within the range observed for the GX-119 materials. The 34% crystallinity content for RAI-116 is significantly lower than the 43% value for GX-119. The lower solubility of RAI-116 (14% versus 21% for GX-119) in good PE solvents reflects the lower crystalline PE content.

Table 1. Grafted–crosslinked polyethylene film RAI-116

Property	Determination
Density at 22°C,[a] g/cm^3	1.218
Melting point,[b] °C	105
Poly (Potassium Acrylate), wt %	32[c]
Polyethylene, wt %	68
Poly (Acrylic Acid), wt %	Undetected
PE crystallinity,[d] %	34
Solubility in good acrylic solvents,[e] %	0
Solubility in good PE solvents,[f] %	14

[a]Determined from buoyancy in silicone oil.
[b]Determined by differential scanning calorimetry (DSC).
[c]Determined from K determination by flame photometry.
[d]Determined from DSC using Petrothane 301 as reference.
[e]Methanol and KOH–Methanol solutions.
[f]α-Chloro-Naphthalene.

9. Role of Crystallinity

The results obtained prior to this study show that the final battery separators have retained a considerable proportion of the initial crystallinity present in the starting PE. This means that all of the crosslinking and grafting has been confined to the amorphous part of the PE initially present. This necessarily leads to an inhomogeneous distribution of the grafted acrylate on the PE as well as some restriction to the amounts of the grafted material.

The battery separator membrane functions to keep separate the constituents of the respective half-cells of the battery, but must permit the ready permeation of the hydroxyl group (the charge carrying species common to both half-cells). As crystallinity in membranes retards permeation, the amount of material transported per unit time for a given membrane decreases noticeably as the crystallinity content increases.

Finally, the dimensional stability of the membrane can be seriously affected since large changes in dimensions are encountered when the very dense crystalline regions melt. Environment, temperature, and mechanical strain all influence the kinetics of recrystallization, and, hence, the rate at which the membrane returns to its former state after heating. If the membrane is first properly fitted for its application and then subsequently heated (as during sterilization), the accompanying dimensional changes may present a serious operational problem, especially if recrystallization proceeds slowly, or not at all, and results in permanent dimensional changes. Evidence for this latter point comes from DSC measurements on heat-sterilized specimens whose crystallinity after one week was about 70% of the initial crystallinity. The role of crystallinity in membrane performance should not be overlooked. The fact that the RAI-116 membrane with its lower crystallinity content improved the performance of test cell operation suggests that crystallinity should be kept at a minimal level.

Reference

1. Adams, L. M., Harlowe, W. W., Jr., and Lawrason, G. C., *Fabrication and Testing of Battery Separator Material from Modified Polyethylene*, Project 01-1842 Final Report, Southwest Research Institute, San Antonio, Texas, June 7, 1966.

C. The Ethylene Oxide–Freon 12 Decontamination Procedure: Control and Determination of the Moisture Content of the Chamber,

R. H. Silver and S. H. Kalfayan

1. Introduction

The rate of decontamination with ethylene oxide (ETO) is dependent, among other factors, upon the moisture content, or relative humidity (RH), in the decontamination chamber. Decontamination is believed to be most effective at RH values of 35–55%.[1]

This article concerns the evaluation of several selected, commercially-available, humidity-sensing and controlling instruments. A literature search indicated that very little work has been done on moisture determination in atmospheres other than air or inert gases (Refs. 1, 2). Available instruments have been developed mainly for the determination of the moisture content in air; however, some appeared promising with respect to the capability of determining the moisture content in an ETO–Freon 12 atmosphere.

To be usable in an ETO–Freon 12 atmosphere, the sensing instrument should fulfill the following requirements:

(1) Its sensing probe should remain unaffected by the ETO–Freon 12 mixture for long periods at 50°C.

(2) It must be capable of remotely reading the moisture content in the decontamination chamber.

(3) It must be accurate and give repeatable results.

(4) Its cost and operation should be inexpensive.

Among the many basic types of humidity-sensing instruments available (Ref. 2), those considered most suitable for evaluation in ETO–Freon 12 were the electrical-resistance type, the electrical-impedance type, and the cold-mirror optical dew-point type.

2. Experimental Instruments

a. Description of instruments evaluated.

Electrical-resistance instrument[2] (Ref. 3). This instrument consists of an electric circuit capable of measuring the resistance of a proprietary wafer made of sulfonated polystyrene. Gold electrodes are deposited on the surface and, as the surface resistance changes due to the wafer adsorbing moisture, the output is read from a meter directly as %RH. The instrument is calibrated periodically by substituting a known resistance in place of the sensor element.

Electrical-impedance instrument[3] (Ref. 4). The operation of this sensor depends upon the impedance of a thin layer of aluminum oxide placed between pure aluminum and gold faces. By means of a suitable measuring circuit, the electrical impedance of the sensor is displayed on a meter, the value of which can be converted to the dew point using a calibration curve supplied by the manufacturer. The basic element, known as the Stover type, is illustrated in Fig. 2. The electrical behavior of this type of element may be described in terms of a single pore of aluminum oxide and an equivalent circuit for that pore.

Cold-mirror optical dew-point instrument[4] (Ref. 5). This instrument utilizes an optical-sensing technique, a cooling element and a linear thermometer (Fig. 3). The optical technique detects the formation of dew on the sensor mirror. The temperature is monitored by the linear thermometer, the output of which is displayed as a direct dew-point reading on a meter. The cooling of the sensor mirror depends upon the bismuth-telluride semiconductor Peltier effect. The more important properties of this semiconductor are expressed by a figure of merit, Z, as

$$Z = f\left(\frac{\pi^2}{T_c^2 \rho K}\right) \tag{1}$$

[1]JPL Specification VOL-50503-ETS.

[2]El-Tronics, Inc., Warren, Pa., Model 102.

[3]Panametrics, Inc., Waltham, Mass., Model 1000.

[4]Technology–Versatronics, Inc., Yellow Springs, Ohio, Model 707.

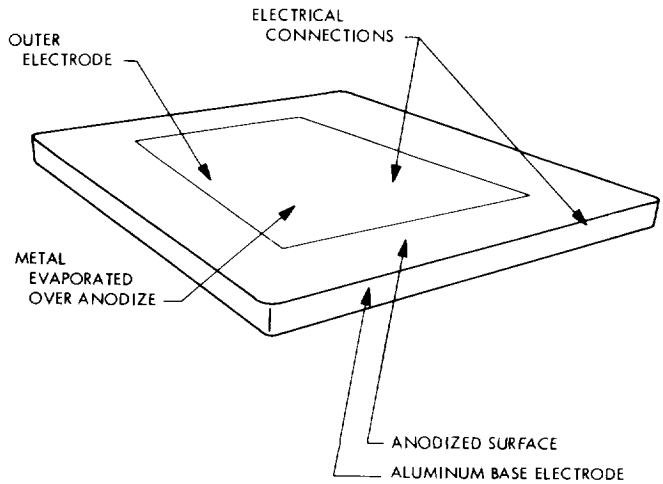


Fig. 2. Stover-type, aluminum-oxide humidity element

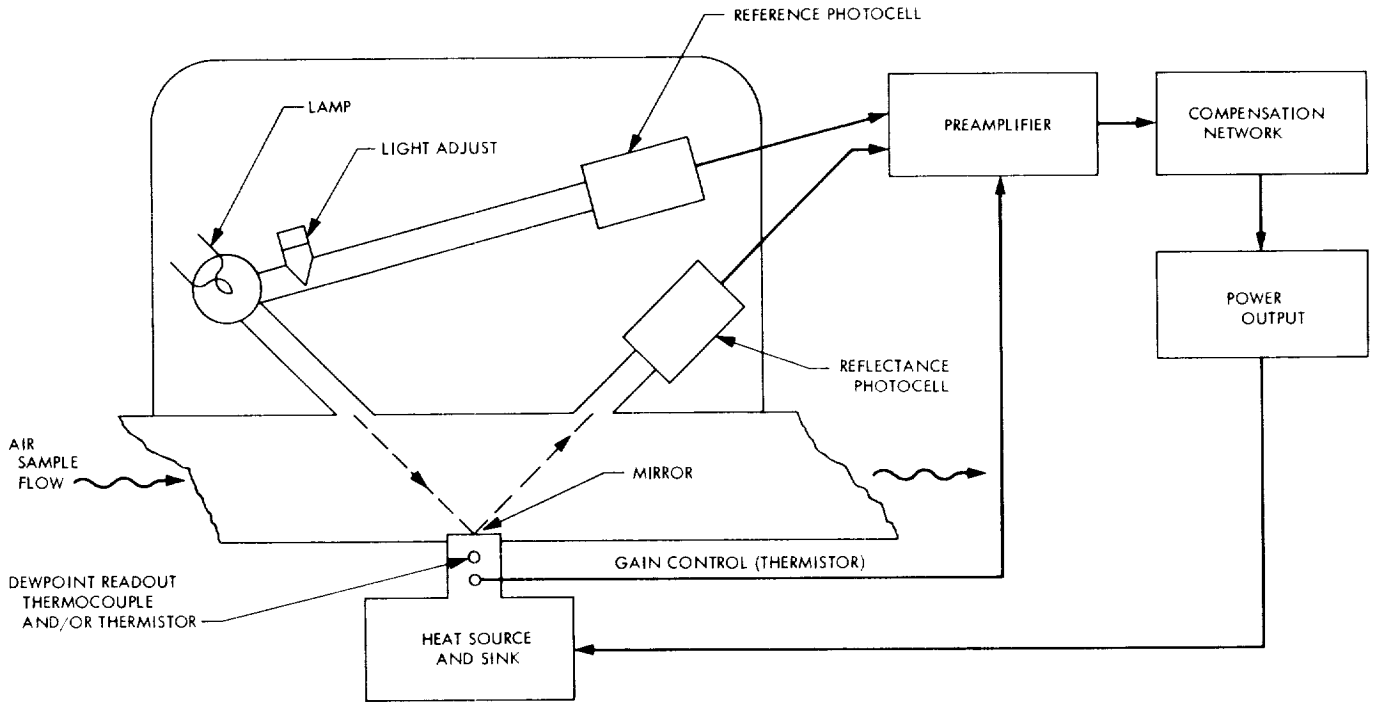


Fig. 3. Functional block diagram for a typical cold-mirror hygrometer

where

π = the Peltier coefficient

K = the thermal conductivity

T_c = the cold-face temperature

ρ = the electrical resistivity

For good cooling, Z should be maximized; therefore, ρ and K should have values as low as possible, and π should be as high as possible. The best Z value currently obtainable from commercial materials is about $3 \times 10^{-3}\ ^{\circ}C^{-1}$.

b. Description of reference instruments. Two other types of humidity-sensing instruments were used in the present investigation. One, a manual dew-point instrument, served as a standard of reference, and the other, a wet-and-dry bulb psychrometer, was used as a secondary reference for comparison of data. As both of these instruments were unsuitable for use in the ETO–Freon atmosphere, they were used as references for measurements taken in air. A brief description of these instruments is given below.

Manual dew-point instrument[5] (Ref. 6). In the operation of this instrument, the gas mixture is drawn into an observation chamber at a pressure above that of the atmosphere and a gauge indicates directly the ratio between the pressure of the gas sample and that of the atmosphere. An operating valve is then depressed, allowing the gas sample in the observation chamber to expand rapidly to atmospheric pressure. In the meantime, a lamp illuminates the chamber, and if the gas has cooled below its dew point (due to its rapid expansion), a distinctive fog is observed in the chamber through a lens system. This procedure is repeated to find the pressure ratio at the time the fog just starts to vanish. The dew point of the gas sample may then be determined by

$$T_{\text{dewpoint}} = (T_i + 460)\,(PR)^{Q} - 460 \qquad (2)$$

where

T_i = initial temperature of the gas sample

PR = pressure ratio

$Q = \dfrac{K-1}{K} = 0.2855$ for air

$K = \dfrac{C_p}{C_v}$

C_p and C_v are specific heats of the gas at constant pressure and constant volume, respectively.

[5]Alnor Instrument Co.—Division of the Illinois Testing Laboratories, Inc., Chicago, Ill.

Wet-and-dry bulb psychrometer[6] (Ref. 7). In this instrument, a battery operated fan causes the ambient air to circulate uniformly over both the wet and dry bulbs at an optimal air velocity for the instrument. The wet bulb is covered with a water-soaked cotton sleeving. Readings of both the wet and dry bulbs are taken after the instrument has reached a state of equilibrium. The averages of several readings are recorded.

c. Test procedure. The procedure followed in testing the instruments is shown diagramatically in Fig. 4. For the first exposure conditions, namely at room temperature in air, readings were taken twice a day for one week with all the instruments. At least five readings were taken with each instrument at the other exposure conditions.

d. Data reduction. With the exception of the electrical-resistance instrument, all instruments gave readings in some parameter other than RH. The data obtained were converted to RH values to facilitate the comparative evaluation.

Raw data from the wet-and-dry bulb psychrometer consist merely of the temperatures of the wet and the dry bulbs. The values of RH, which are obtainable from psychrometric tables (Ref. 8), are based on the Ferrel equation

$$e = e' - 0.000367P(t - t')\left(1 + \frac{t' - 32}{1571}\right) \qquad (3)$$

where t is the temperature of the dry bulb, t' is the temperature of the wet bulb, P is the barometric pressure of air in inches, e' is the saturation pressure of water vapor at t'. Solving the Ferrel equation gives e, the vapor pressure of water at t.

Readings from the manual dew-point instrument consist of the pressure ratio and the ambient temperature. The dew point can be calculated by the use of Eq. (3); however, the RH can be obtained from the dew point using the relationship

$$\frac{\text{vapor pressure at dew point}}{\text{vapor pressure at room temp}} \times 100 = \%\text{RH}$$

The use of the Smithsonian Tables (Ref. 9) facilitates the data reduction.

[6]Bendix Instruments, Inc., Model 566.

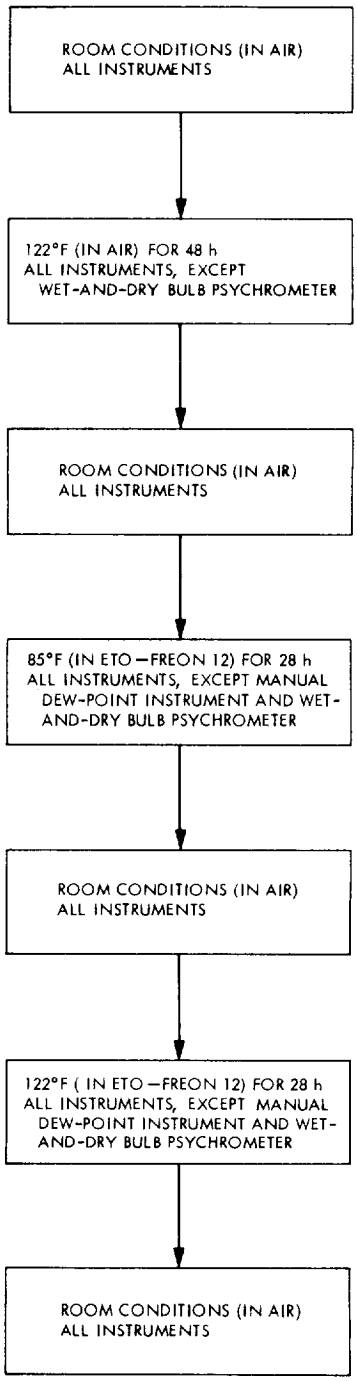


Fig. 4. Test sequence for the evaluation of moisture detectors

3. Results and Discussion

The initial RH readings in air obtained from the manual dew-point instrument and the test instruments differed significantly in that variations of up to 4 percentage points in RH were observed. The disagreements between the test

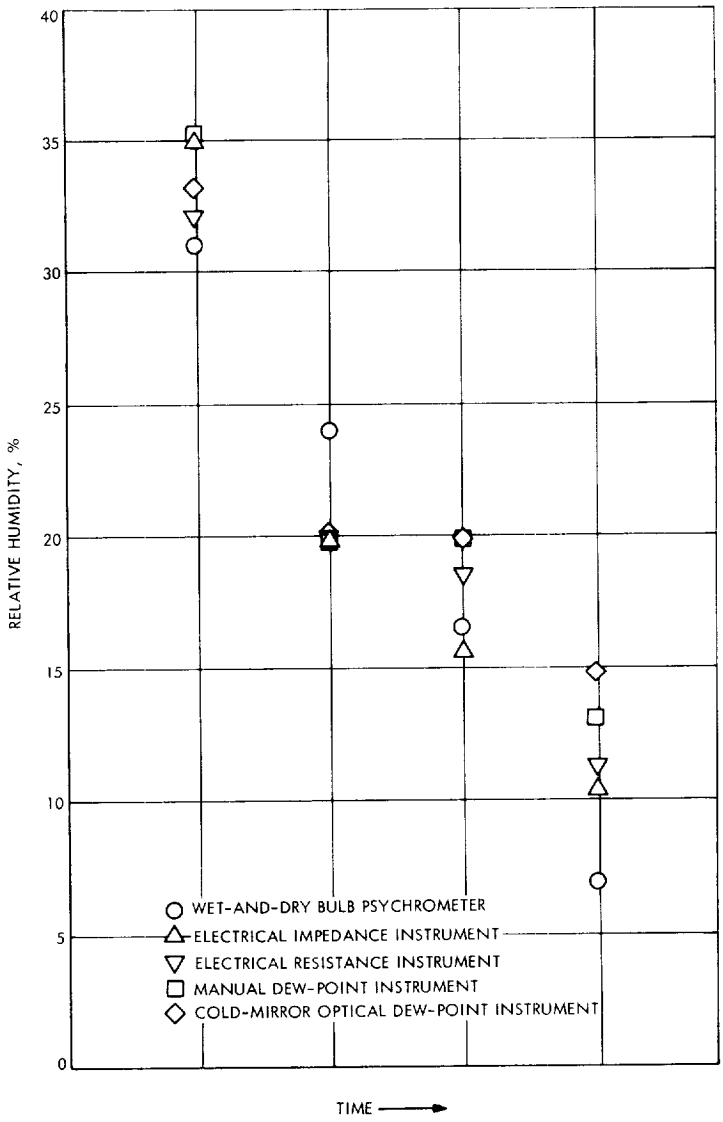


Fig. 5. Readings for various humidity-sensing instruments at room conditions

instruments themselves were also significant. Figure 5 shows these disagreements plus the values obtained by the wet-and-dry bulb psychrometer and the manual dew-point instrument.

The tendency of the cold-mirror optical dew-point instrument, compared to the manual dew-point instrument, was to read higher RH values, while the tendency of the electrical-impedance and electrical-resistance instruments was to indicate lower RH values, also as compared to the manual dew-point instrument.

Significant changes occurred in the probes of the electrical-impedance and cold-mirror optical dew-point instruments when they were exposed to ETO–Freon 12 either at room temperature or at 50°C. Under these conditions, readings from the electrical-impedance instrument probe were exceedingly low, while those from the cold-mirror optical dew-point instrument probe were quite high. Bringing the probes to atmospheric conditions did not improve the situation.

A vacuum treatment of the electrical-impedance instrument probe, to dissipate any absorbed sterilant gas, did not improve its operation. Likewise, repeated cleaning of the mirror of the cold-mirror optical dew-point instrument did not bring about a recovery. Indications were that the probes of both instruments were irreversibly affected

by the ETO–Freon 12 gas mixture. Presumably, the aluminum-oxide of the electrical-impedance instrument probe and the bismuth-telluride semiconductor of the cold-mirror optical dew-point instrument probe reacted with the ETO–Freon 12 atmosphere to change their capacitance or resistance.

The electrical-resistance instrument probe was the least affected; however, readings taken from it after one cycle of ETO–Freon 12 exposure were much lower, with reference to the manual dew-point instrument, than those taken prior to exposure. Before exposure, the %RH readings were on the order of 1–3 percentage points below those of the manual dew-point instrument; after exposure, these readings dropped to 8–10 percentage points below those of the reference instrument. While washing the sensor element of the electrical-resistance instrument with distilled water improved its sensitivity, further exposure to ETO–Freon 12 again lowered its sensitivity. There was, however, no indication of a continuous deterioration of the element as it was exposed to subsequent cycles of ETO–Freon 12 decontamination. Additional washing with distilled water did not improve the probe's response.

4. Conclusion

The humidity sensors evaluated thus far have not proved to be suitable for moisture determination in an atmosphere of ETO–Freon 12. Although one of these, the electrical-resistance type sensor, appears to be far better than the others, it tends to read lower %RH values after exposure to ETO–Freon 12. This study has confirmed the need for humidity sensors that will function properly in an atmosphere of this sterilant gas mixture.

References

1. Flanigan, F. M., "Comparison of the Accuracy of Humidity Measuring Instruments," *J. ASHRAE*, Vol. 2, pp. 56–59, Dec. 1960.
2. *Humidity and Moisture; Measurement and Control in Science and Industry*: Vols. I–III. Edited by A. Wexler, Reinhold Publishing Co., New York, 1965.
3. Technical Bulletin 4, TB-31-66-K, El-Tronics, Inc., Warren Components Division, Warren, Pa.
4. *Aluminum-Oxide Hygrometer Model 1000*, Specification and Operating Manual, Panametrics, Inc., Waltham, Mass., 1968.
5. Paine, L. C., and Farrah, H. R., "Design and Applications of High-Performance Dew-point Hygrometers," in *Humidity and Moisture: Vol. I.*, pp. 174–188, Edited by A. Wexler. Reinhold Publishing Co., New York, 1965.
6. *Alnor Dewpointer Instructions*, Illinois Testing Laboratories, Inc., Bulletin 72-2051, Sept. 1966, and Form 7240, Oct., 1959.
7. *Wet-and-Dry Bulb Psychrometer Method*, Technical Bulletin 1, Hygrodynamics, Inc., Silver Springs, Maryland.
8. Marvin, C. F., *Psychrometric Tables*, WB 235, p. 9. U. S. Department of Commerce, Weather Bureau, Washington, D. C., 1941.
9. *Smithsonian Meteorological Tables*, pp. 350–359, 6th Revised Edition. Smithsonian Institute, Washington, D. C., 1951.

XI. Research and Advanced Concepts

PROPULSION DIVISION

A. Special Applications for Spectroscopic Scanning of Internal Plasma Flows, *E. J. Roschke*

1. Introduction

Radial distributions of temperature and/or electron density in an axisymmetric plasma are obtained commonly from edge-on spectroscopic observations transformed to yield radial distributions of emission. When it is difficult or inconvenient to make edge-on or transverse observations across the whole plasma, one observation location (for example, along a single diameter of the plasma) must suffice. The interpretation of the weighted temperature determined at a single observation location was examined in SPS 37-47, Vol. III, pp. 116–119, for slightly ionized argon, assumed to be optically thin and having an axisymmetric temperature distribution with a peak at the duct centerline.

The purpose of this article is to enlarge upon the previous results and extend the discussion to the more general case of off-axis observations. In particular, it became desirable to determine if inversion of edge-on data (by the Abel formula) was necessary for the usual conditions of the experiments. This has practical significance aside from accuracy since the standard inversion techniques require axial symmetry whereas the actual plasma flow may not be axisymmetric. Even though the results given later apply to symmetric flows only, the general conclusions will not be seriously modified by the presence of moderate degrees of asymmetry, provided absorption in the gas is absent.

An optically thin (nonabsorbing), slightly ionized gas is considered; the temperature distribution is assumed to be axisymmetric, having a peak value at the axis of the flow. In particular, interest is confined to argon in thermal equilibrium at temperatures less than 10^{4}°K and pressures of the order of 0.1 to 0.3 atm. The latter considerations eliminate the possibility of off-axis peaks in the radial intensity distributions of argon atom lines and reduce the emission of ion lines to virtually zero. Thus, the results are not general and must be considered only in terms of the assumptions.

2. Viewing Along a Diameter—Further Remarks

The general expression for the spectral line-intensity is given as (Ref. 1)

$$I = \frac{gAC}{\lambda} \frac{N}{u} e^{-\frac{E}{kT}} \tag{1}$$

(See Table 1 for definition of terms.)

Table 1. Nomenclature

A	transition probability
B	value of integral, e.g., Eq. (8)
b	ratio of temperature at center of duct to wall temperature, i.e., T_0/T_w
C, C', C''	constants in intensity equation, independent of λ or E
D	diameter of duct
E	upper energy level of excited states
g	statistical weight
I	intensity of radiation of a spectral line
$I(x)$	total intensity of a line as seen by spectrometer viewing along a chord or diameter of plasma
K	constant view factor independent of λ or X
k	Boltzmann's constant
m	parameter defined by E/kT_0 or $E/kT_0(x)$
N	number density of emitting species
n	exponent defined in text
R	radius of duct
r	radial coordinate
$\bar{T}$	temperature, absolute
$T(x)$	temperature along x-axis
T	weighted temperature determined spectroscopically by viewing along a chord or diameter of plasma
u	partition function
x, y	coordinates transverse and parallel to direction of observation, respectively (Fig. 1)
(x)	function of x, used with B, I, T, and m
X	distance from y-axis of an off-axis observation (Fig. 1)
Y	half-chord length corresponding to X (Fig. 1)
z	axial coordinate
α	ionization fraction
β	coefficient equal to $(b-1)/b$
λ	wavelength
ϕ	coefficient defined in text
Subscripts	
1, 2	designations for parameters corresponding to wavelengths λ_1 and λ_2
0	value at center of duct, or along x-axis when used with (x)
w	value at boundary or wall
*	indication for normalized value, defined as used in text

If two experimental measurements of intensity $I_1(x)$ and $I_2(x)$ are made along a diameter of the plasma corresponding to the spectral lines λ_1 and λ_2, then the weighted temperature for the path of view according to SPS 37-47, Vol. III, may be calculated from

$$\bar{T} = \frac{(E_1 - E_2)\log e}{k \log\left[\frac{\lambda_2 I_2(x)}{g_2 A_2 C_2} \cdot \frac{g_1 A_1 C_1}{\lambda_1 I_1(x)}\right]} \tag{2}$$

Equation (2) is strictly correct only when the number density of the emitting species has a weighted value, over the path of view, which bears a simple relationship to $\bar{T}$, e.g., when $\bar{N}$ is proportional to the reciprocal of $\bar{T}$. The weighted temperature $\bar{T}$ is not a spatial average, as pointed out in Ref. 2.

An integrated value of $I(x)$ across a diameter of the plasma is calculated from Eq. (8), p. 117, of SPS 37-47, Vol. III, as

$$\frac{\lambda I(x)}{gAC'} = \frac{DB}{T_0} e^{-m} \tag{3}$$

where $m = E/kT_0$ and B depends on the shape of the temperature distribution as well as E. In general, it appears that for temperature distributions which decrease monotonically with radius, $B = \text{const} \cdot E^{-n}$ with $n \leq 1$. Substituting Eq. (3) into Eq. (1), taking logs and rearranging,

$$\log\left[\frac{\lambda E^n I(x)}{gAC''}\right] = -\frac{\log e}{kT_0} E \tag{4}$$

where C'' is a new constant independent of λ or E. Hence, this formulation suggests that the centerline temperature T_0 should be obtained from a relative line-intensity method viewing across a diameter by including the term E^n in the dependent variable. Unfortunately, n is not generally known, except $n = 0$ for a uniformly flat temperature distribution. For a parabolic temperature distribution, $n = 1/2$ (SPS 37-47, Vol. III).

3. Off-axis Viewing When Temperature Distribution is Parabolic

Enthalpy distributions across a circular duct using arc-heated argon have been obtained by means of a calorimetric probe (Ref. 3). Generally, the results indicate the existence of a flat, adiabatic core surrounded by a cool, annular boundary layer. The boundary layer grows progressively thicker with increasing distance from the inlet until eventually it occupies the entire duct. The case of an off-axis direction of view (as viewed by the spectrometer) will now be considered in a manner similar to that used for diametral viewing in SPS 37-47, Vol. III. An integrated spectral intensity along an arbitrary chord will be obtained for the case of an assumed parabolic temperature distribution for purposes of illustration. The weighted temperature as obtained spectroscopically will be compared to the maximum temperature existing in the path of view.

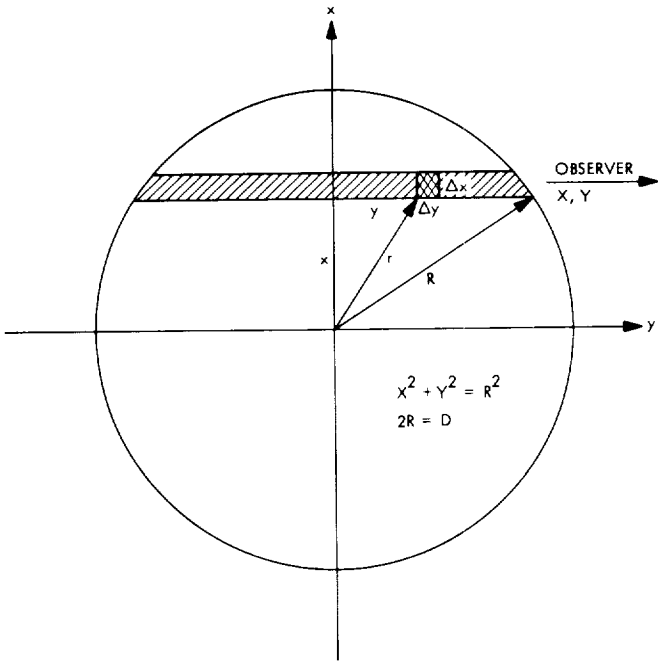


Fig. 1. Cross section of axisymmetric plasma

Notation is given in Fig. 1. The diameter of the duct D is coincident with the y-axis; Y represents the half-chord length corresponding to an off-axis direction of view parallel to the y-axis at a distance X from the y-axis. The maximum temperature in the field of view is denoted by $T_0(x)$ and occurs at the x-axis. T_0 is the maximum temperature at the center of the duct, i.e., at $x = y = 0$. A parabolic temperature profile in the plasma is given by

$$T_* = \left(\frac{T}{T_0}\right) = 1 - \beta r_*^2 \tag{5}$$

where $\beta = (b - 1)/b$, $b = T_0/T_w$, and $r_* = r/R$. Along the direction of view, Eq. (5) can be transformed to the following relationship:

$$T_*(x) = \frac{T(x)}{T_0(x)} = 1 - \phi\left(\frac{y}{Y}\right)^2 \tag{6}$$

where

$$\phi = \frac{\beta Y_*^2}{1 - \beta X_*^2}$$

$$X_* = \frac{X}{R}$$

$$Y_* = \frac{Y}{R}$$

The integrated spectral intensity in the direction of view for a slice dx high by dz deep is

$$\frac{\lambda I(x)}{gAC} = 2K\int_0^Y \frac{N}{u} e^{-E/kT}\, dy \tag{7}$$

where K is small compared to unity. Assuming N is proportional to $1/T$ and $u \sim 1$, with $T_*(x)$ given by Eq. (6), Eq. (7) in normalized form becomes

$$\frac{\lambda I(x)}{gAC'} = \frac{2Ye^{m(x)}}{T_0(x)} \times \int_0^1 \left[\frac{1}{T_*(x)}\right] \exp\left\{-m(x)\left[\frac{1}{T_*(x)} - 1\right]\right\} d\left(\frac{y}{Y}\right) \tag{8}$$

where $m(x) = E/kT_0(x)$. Equation (8) is identical in form to Eq. (8) of SPS 37-47, Vol. III, p. 118. Hence, the value of the integral $B(x)$ is given by numerical integration as

$$B(x) = \text{const} \cdot [\phi m(x)]^{-1/2} \tag{9}$$

For a given value of X, ϕ will be the same for all spectral lines so that on a relative basis $B(x)$ varies only as $m(x)$ varies. Thus, if two spectral lines at λ_1 and λ_2 are used to determine the weighted temperature along the direction of view,

$$\frac{\bar{T}(x)}{T_0(x)} = \frac{[m_1(x) - m_2(x)] \log e}{[m_1(x) - m_2(x)] \log e + \log \left[\frac{B_2(x)}{B_1(x)}\right]} \tag{10}$$

where

$$\frac{B_2(x)}{B_1(x)} = \left[\frac{m_1(x)}{m_2(x)}\right]^{1/2} = \left(\frac{E_1}{E_2}\right)^{1/2}$$

Equation (10) is analogous to Eq. (9) of SPS 37-47, Vol. III, and reduces to it when $X = 0$, $Y = R$. Considering argon atom lines at values of temperature 10^4°K or less, $\bar{T}(x)$ will differ from $T_0(x)$ by less than 5%.

From Eq. (10), it is clear that the largest relative discrepancy in temperature occurs at the y-axis (diameter) and decreases as X increases. Also, the discrepancy increases with increasing T_0 (center temperature) and increasing b (ratio of center temperature to wall temperature). In general then, the weighted temperature obtained along a chord will agree with the maximum temperature along that direction of view to within a few percent. Under these conditions, Abel inversion of experimental, off-axis, intensity measurements is not necessary; and the temperature distribution obtained from these measurements may be assumed to be the radial distribution if symmetry is present. However, if the actual enthalpy and temperature distributions are not axisymmetric, only the maximum temperature in parallel slabs of plasma is obtained. Whether or not symmetry is present, the temperature determined spectroscopically will agree with the maximum temperature in the path of view with increasing accuracy accordingly as the relative portion of flow having a flat temperature distribution increases in extent.

In the case of internal flows, a problem exists which seems to have been discussed rarely in the literature. This concerns the possible errors introduced by a highly reflective internal wall. If this wall were a perfect reflector over all wave lengths, the plasma emission would be uniform over the volume and the integrated intensity distribution would be elliptical (Ref. 4). The degree of error introduced will depend on the actual reflectivity of the wall and whether it is a diffuse or specular reflector. In the present experiments, stainless steel parts are usually used; and it is possible that these have sufficient normal and hemispherical reflectivity to introduce some error, especially in the cool regions of the plasma. However, as pointed out in Ref. 4, the problem can be avoided or greatly reduced if observations are made against the background of a deep cavity, thus providing an essentially black background for observation. Such is the case in the present experiments.

4. Conclusions

For axisymmetric and optically thin internal flows of slightly ionized argon at specified conditions, it has been shown that temperatures determined spectroscopically from off-axis or edge-on views tend to agree closely with the maximum temperature in the path of view. The agreement is poorest at the center of the duct and also tends to decrease with an overall increase in temperature level. It is concluded that Abel inversion of off-axis intensity data to obtain radial intensity distributions is not necessary when the conditions of the present treatment are met. Thus, the transverse temperature distribution obtained from scanning may be taken as the true radial distribution of temperature with very little error. These results are not general, however, and do not necessarily apply when the center temperature greatly exceeds 10^4°K or when the ionization fraction α begins to approach 0.1 or more.

References

1. Pearce, W. J., *Optical Spectrometric Measurements of High Temperatures*, pp. 125–169. Edited by P. J. Dickerman. University of Chicago Press, Chicago, Ill., 1960.
2. Oertel, G. K., *Remarks on Practical Spectroscopic Temperature Measurements in Plasmas*, NASA TN D-3737. National Aeronautics and Space Administration, Washington, D.C., Dec. 1966.
3. Massier, P. F., Back, L. H., and Roschke, E. J., "Heat Transfer and Laminar Boundary-Layer Distributions in an Internal Subsonic Gas Stream at Temperatures Up to 13,900 Deg R," Paper No. 68-HT-16, presented at AIChE-ASME Heat Transfer Conference, Philadelphia, Pa., Aug. 1968.
4. Zaidel, A. N., Malyshev, G. M., and Shreider, E. Y., "Spectroscopic Diagnostic Techniques for Hot Plasmas," *Sov. Phys.—Tech. Phys.*, Vol. 6, No. 2, pp. 93–119, Aug. 1961.

B. Hall and Ion-Slip Effects in Channel Flow,

E. J. Roschke

1. Introduction

The purpose of this article is to consider both Hall and ion-slip effects on the joule-heating parameter S, which influences the convective heat transfer from an

electrically conducting fluid for laminar flow. Calculated results for argon are presented so that the relative magnitudes of both the Hall and ion-slip parameters may be compared and their effect on S evaluated. (See Table 2 for definition of terms.)

Back[1] considered laminar heat transfer for constant property slug flow between infinite parallel plates subject to an applied, uniform, transverse magnetic field. Provision for an applied electric field was also included. The effect of an applied magnetic field arises by joule heating, which appears in the energy equation in the term j^2/σ. This term, when nondimensionalized, is called the joule-heating parameter S, which may be written as $S = Ha^2Ek(1 - K^2)Pr$. The influence of the Hall effect on S was considered briefly in SPS 37-47, Vol. III, pp. 121–128. It was found that j^2/σ or S was reduced by a factor $[1 + (\omega_e\tau_e)^2]$, where $\beta_e = \omega_e\tau_e$ is called the Hall parameter. This result may be viewed either as a decrease in current or, more usually, as a decrease in electrical conductivity. For temperatures in the range 7,000 to 14,000°K and pressures in the range 0.1 to 1.0 atm, it was found that β_e decreases with increasing temperature and increasing pressure and can take on values considerably greater than unity. When $\beta_e \gg 1$, a drastic decrease in S results theoretically.

2. General Considerations

The collision frequencies are defined by

$$\omega_e = eB/m_e \text{ and } \omega_i = eB/m_i$$

and the times between collisions are τ_e and τ_i; the Hall and ion-slip parameters are defined by $\beta_e = \omega_e\tau_e$ and $\beta_i = \omega_i\tau_i$. Neglecting electron pressure gradient, the generalized form of Ohm's law (Ref. 1) may be expressed in the following form:

Table 2. Nomenclature

B	magnetic field vector (applied)
E	electric field vector (applied)
Ek	Eckert number
e	charge on electron
Ha	Hartmann number
j	current density
j	current density vector
K	load factor, ratio of applied to induced electric fields
k	Boltzmann's constant
m	particle mass
n	particle number density
Pr	Prandtl number
q	general velocity vector
q	collision cross section
S	joule-heating parameter
S′	modified joule-heating parameter
T	temperature, absolute
u, v, w	velocity components (Fig. 2)
x, y, z	spatial coordinates (Fig. 2)
β_e, β_i	Hall and ion-slip parameters, respectively
σ	scalar electrical conductivity
τ	time between particle collisions
ω	collision frequency
Subscripts	
a, e, i	atom, electron, and ion, respectively
x, y, z	designate components along coordinate axes

$$\mathbf{j} = \frac{\sigma}{(1 + \beta_e\beta_i)^2 + \beta_e^2}\left\{(1 + \beta_e\beta_i)(\mathbf{E} + \mathbf{q} \times \mathbf{B}) - \beta_e(\mathbf{E} + \mathbf{q} \times \mathbf{B}) \times \frac{\mathbf{B}}{|B|}\right.$$
$$\left. + [\beta_e^2 + \beta_i(1 + \beta_e\beta_i)]\left[(\mathbf{E} + \mathbf{q} \times \mathbf{B}) \cdot \frac{\mathbf{B}}{|B|}\right]\frac{\mathbf{B}}{|B|}\right\} \quad (1)$$

If there is no component of current in the direction of **B**, then the third term on the right side vanishes. Ion-slip is included in the first term and produces an additional transverse current as that arising simply from $\mathbf{q} \times \mathbf{B}$. The Hall current arises from the second term and is in the direction of flow if the flow is purely axial and no axial component of applied **E** exists.

[1]Back, L. H., *Laminar Heat Transfer in Electrically Conducting Fluids Flowing Between Parallel Plates* (submitted for publication in *Int. J. Heat and Mass Transfer*).

It is now assumed that (Fig. 2)

$$\mathbf{j} = (j_x, 0, j_z)$$
$$\mathbf{E} = (E_x, 0, E_z)$$
$$\mathbf{q} = (u, 0, w)$$
$$\mathbf{B} = (0, B_y, 0)$$

From Eq. (1), the expressions for the axial and transverse components of current become

$$j_x = \frac{\sigma}{(1 + \beta_e\beta_i)^2 + \beta_e^2} [(1 + \beta_e\beta_i)(E_x - wB_y) + \beta_e(E_z + uB_y)] \tag{2}$$

$$j_z = \frac{\sigma}{(1 + \beta_e\beta_i)^2 + \beta_e^2} [(1 + \beta_e\beta_i)(E_z + uB_y) - \beta_e(E_x - wB_y)] \tag{3}$$

When properly modified, these expressions agree with those given in Refs. 2 and 3. From Eqs. (2) and (3), the joule-heating term becomes

$$\frac{j^2}{\sigma} = \frac{j_x^2 + j_z^2}{\sigma} = \frac{\sigma}{(1 + \beta_e\beta_i)^2 + \beta_e^2} [(E_x - wB_z)^2 + (E_z + uB_y)^2] \tag{4}$$

Table 3. Current components and joule-heating for $E_x = w = 0$

Case	Axial current j_x	Transverse current j_z	Joule heating j^2/σ
$\beta_i = 0$	$\frac{\sigma\beta_e}{1+\beta_e^2}(E_z + uB_y)$	$\frac{\sigma}{1+\beta_e^2}(E_z + uB_y)$	$\frac{\sigma}{1+\beta_e^2}(E_z + uB_y)^2$
$\beta_e = 0$	0	$\sigma(E_z + uB_y)$	$\sigma(E_z + uB_y)^2$
$\beta_e = \beta_i = 0$	0	$\sigma(E_z + uB_y)$	$\sigma(E_z + uB_y)^2$

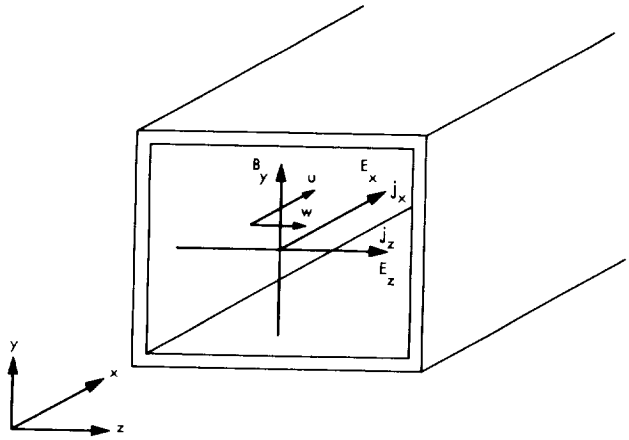


Fig. 2. Geometry and coordinate system for channel flow

The effects of the Hall and ion-slip parameters are listed in Table 3 for the case of $E_x = 0$ (continuous conductor) and $w = 0$ (zero cross-velocity). It is clear from Eqs. (2) and (3) that ion-slip does not increase the transverse current j_z but tends to decrease it; however, the contribution of the Hall current to the total axial current j_x is also decreased by ion-slip. From Eq. (4), it is determined that the reduction in the joule-heating parameter S due to Hall and ion-slip effects is

$$S' = \frac{S}{(1 + \beta_e\beta_i)^2 + \beta_e^2} \tag{5}$$

Thus, it is sufficient to compare the magnitude of β_e with $(1 + \beta_e\beta_i)$ to assess the relative importance of ion-slip.

If $E_x = w = 0$, Eqs. (2) and (3) reduce to

$$j_x = \frac{\sigma\beta_e(E_z + uB_y)}{(1 + \beta_e\beta_i)^2 + \beta_e^2} \tag{6}$$

$$j_z = \frac{\sigma(1 + \beta_e\beta_i)(E_z + uB_y)}{(1 + \beta_e\beta_i)^2 + \beta_e^2} \tag{7}$$

and

$$\frac{j_x}{j_z} = \frac{\beta_e}{1 + \beta_e\beta_i} = \frac{\omega_e\tau_e}{1 + \omega_e\tau_e\omega_i\tau_i} \tag{8}$$

Equation (8) is plotted in Fig. 3 versus β_e for various values of β_i/β_e. Later, it will be shown that $\beta_i \sim 10^{-2}\,\beta_e$ for singly ionized argon in thermal equilibrium in the range of interest. Hence, for $1 < \beta_e < 10$, $j_x > j_z$; and for $\beta_e \leq 1$, $j_x < j_z$. Also, under these conditions, it is clear that ion-slip has a negligible effect on S.

3. Calculation of β_e and β_i and Discussion

Expressions for $\beta_e = \omega_e\tau_e$ and $\beta_i = \omega_i\tau_i$ given in Ref. 2 are

$$\frac{\omega_e\tau_e}{B} \simeq \frac{3}{4}\, e \left(\frac{\pi}{8m_e kT_e}\right)^{1/2} \left(\frac{n_e q_{ei}}{2} + n_a q_{ea}\right)^{-1} \qquad (9)$$

$$\frac{\omega_i\tau_i}{B} \simeq \frac{3}{4}\, e \left(\frac{\pi}{8m_i kT_i}\right)^{1/2} \frac{n_a}{(n_a + n_e)^2\, q_{ia}} \qquad (10)$$

Calculations for singly ionized argon (assuming $T_e = T_i$ and $n_e = n_i$) were made using collision cross sections from Ref. 4; q_{ea} was used directly, q_{ia} was computed from the empirical relation $q_{ia} = (1.44\ T^{0.16})\ q_{aa}$, using q_{aa} from Ref. 4, and q_{ei} (a weak function of pressure) was obtained by extrapolating the low-pressure values given there.

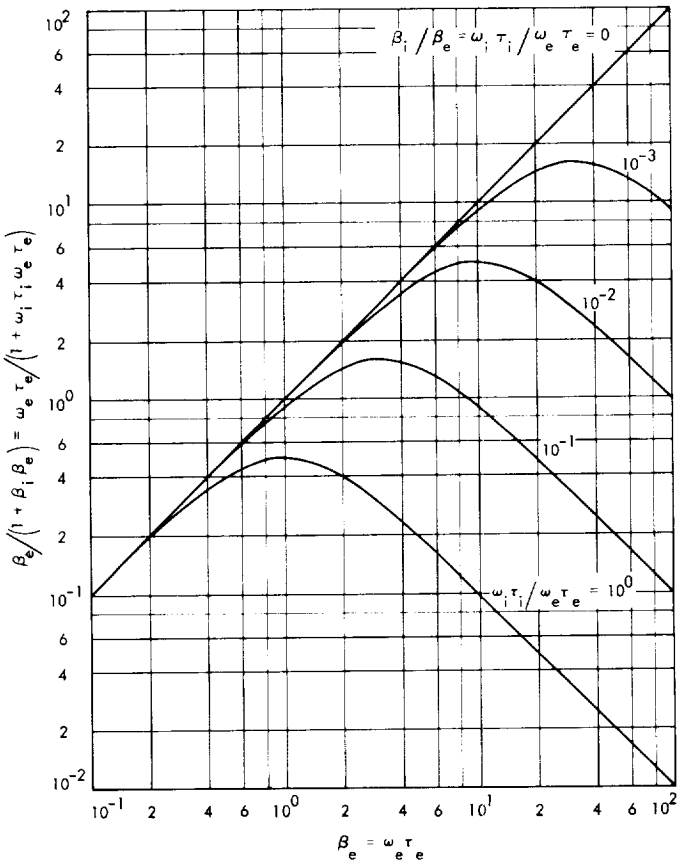


Fig. 3. Ratio of current densities $j_x/j_z = \beta_e/(1 + \beta_i\beta_e)$ for case $E_x = w = 0$

Results are shown in Fig. 4, from which it is seen that $\beta_i \leq 10^{-2}\,\beta_e$. The results given for β_e agree reasonably well with those given in SPS 37-47, Vol. III, which were computed from the simple relation $\beta_e = \sigma B/en_e$. If $\beta_e > 2$, it is likely that T_e is appreciably larger than T_i so that β_i may be somewhat larger relative to β_e than indicated; however, it does not appear likely that ion-slip would contribute significantly to a reduction in joule heating. Theoretically, the Hall effect could affect joule heating significantly even if β_e is only of the order of unity. Experimentally determined values of β_e in the range 1 to 5 are common, but they always tend to be lower than the theoretical values, especially at high values of applied magnetic field (e.g., Ref. 5).

4. Conclusion

Theoretically, ion-slip appears to exert negligible influence on joule heating for argon in the range of interest. Hall effect must be taken into account, however, even though its true magnitude may be considerably less than theoretical. It is sufficiently accurate, in the first approximation, to reduce the joule-heating parameter accordingly as $S' = S/(1 + \beta_e^2)$ with β_e determined experimentally or estimated from the literature. On a theoretical basis, it was found that the simple relation $\beta_e = \sigma B/en_e$ is sufficiently accurate compared to a longer expression utilizing collision cross sections, when suitable values of σ are obtained from results published in the literature.

References

1. Kemp, N. H., and Petschek, H. E., "Two-Dimensional Incompressible Magneto-Hydrodynamic Flow Across an Elliptical Solenoid," *J. Fluid Mech.*, Vol. 4, pp. 553–584, Nov. 1958.
2. Denison, M. R., and Ziemer, R. W., *Investigation of the Phenomena in Crossed-Field Plasma Accelerators,* AIAA Paper 63-378, presented at Fifth Biennial Gas Dynamics Symposium, Northwestern University, Evanston, Ill., Aug. 1963.
3. Kontaratos, A. N., and Demetriades, S. T., "Interaction of a Jet of Plasma with Electric and Magnetic Fields," *Appl. Sci. Res.*, Vol. 11, Sec. B, pp. 335–360, 1964.
4. Cann, G. L., Ziemer, R. W., and Marlotte, G. L., *The Hall Current Plasma Accelerator,* AIAA Paper 63-011, presented at AIAA Electric Propulsion Conference, Colorado Springs, Colo., Mar. 1963.
5. Brederlow, G., and Hodgson, R. T., "Electrical Conductivity of Seeded Noble Gases in Crossed Electric and Magnetic Fields," *AIAA J.*, Vol. 6, No. 7, pp. 1277–1284, July 1968.

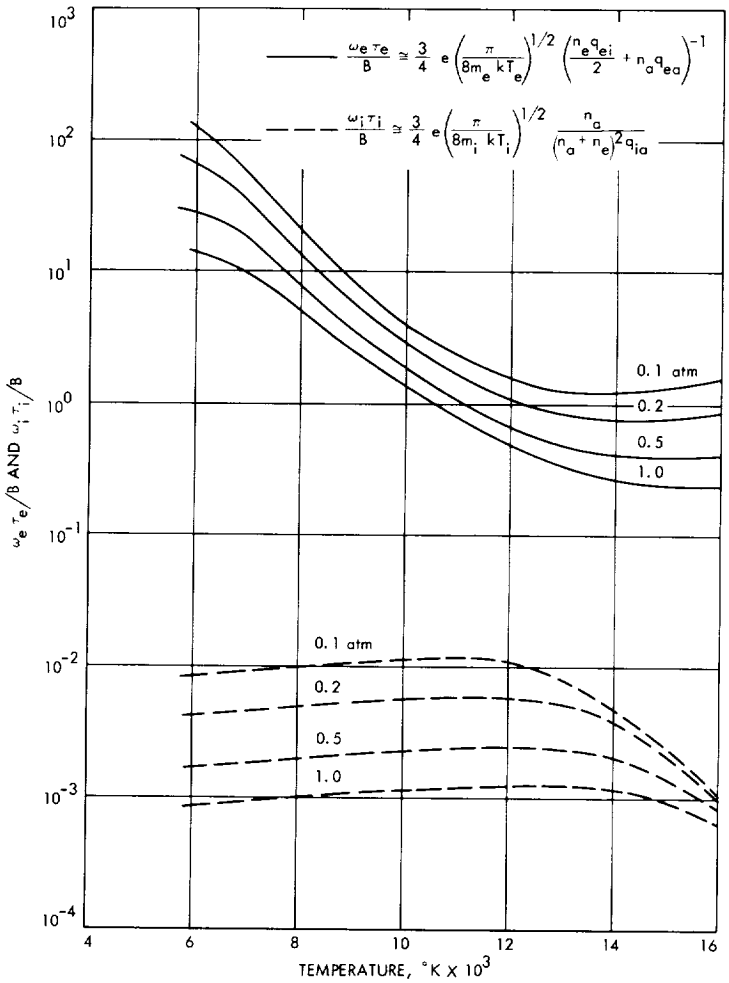


Fig. 4. Hall and ion-slip parameters for singly ionized equilibrium argon

C. Liquid-Metal MHD Power Conversion,

D. G. Elliott and L. G. Hays

1. Introduction

Liquid-metal magnetohydrodynamic (MHD) power conversion is being investigated as a power source for nuclear-electric propulsion. A liquid-metal MHD system has no moving mechanical parts and operates at heat-source temperatures between 1600 and 2000°F. Thus, the system has the potential of high reliability and long lifetime using readily available containment materials such as Nb-1%Zr.

In the particular MHD cycle being investigated, liquid lithium would be heated at about 150 psia in the reactor or reactor-loop heat exchanger; mixed with liquid cesium at the inlet of a two-phase nozzle, causing the cesium to vaporize; accelerated by the cesium to about 500 ft/s at 15 psia; separated from the cesium; decelerated in an alternating-current MHD generator; and returned through a diffuser to the heat source. The cesium would be condensed in a radiator or radiator-loop heat exchanger and returned to the nozzle by an MHD pump.

The ac generator for the NaK-nitrogen conversion system is undergoing empty-channel electrical tests. A theory has been derived for the effect of finite number of slots on MHD induction generator efficiency. Evaluation of the Haynes-Stellite alloy 25 loop with 1800°F lithium was completed, and additional cycle efficiency results were calculated for multistage cycles.

2. Effect of Finite Number of Slots on MHD Induction Generator Efficiency

The ideal MHD induction generator has a continuous winding current sheet $I' = I'_m \cos [(2\pi x/L) - \omega t]$, where

I' is the instantaneous winding current per unit length, I'_m is the amplitude of the winding current, x is the distance from the generator inlet, L is the generator length, and ω is the angular frequency. The current I' produces an empty-channel magnetic field $B_0 = B_{0_m} \sin [(2\pi x/L) - \omega t]$ having amplitude $B_{0_m} = \mu_0 I'_m L/2\pi g$, where g is the iron gap.

When a fluid of conductivity σ flows in the generator through channel height $b \leq g$ and width c, the field is reduced in amplitude and shifted downstream such that

$$B(x,t) = \frac{\mu_0 I'_m L}{2\pi g \left[1 + \left(\frac{sR_m b}{g}\right)^2\right]^{1/2}} \times \sin\left[\frac{2\pi x}{L} - \tan^{-1}\left(\frac{sR_m b}{g}\right) - \omega t\right] \quad (1)$$

where s is the slip $(U - U_s)/U_s$ between the fluid and wave velocities and R_m is the magnetic Reynolds number $\mu_0 \sigma U_s L/2\pi$. The field induces an rms fluid current $I'_f = \sigma b B U_s s$ per unit length, where B is the rms field. The products $cUBI'_f$ and $c(I'_f)^2/\sigma b$ integrated over the generator length give the input power P_m and the ohmic loss P_r, respectively, from which the internal electrical efficiency $\eta_0 = (P_m - P_r)/P_m$ is found to be $(1 + s)^{-1}$. The question for a practical generator is (1) how many slots must the actual generator have to give a sufficiently sinusoidal $I'_m(x)$ for negligible departure from the $(1 + s)^{-1}$ efficiency, and (2) how rapidly does the efficiency diminish with reduction in the number of slots?

The above equations for B and I'_f already contain the solution since the actual winding current I' can be written as a sum of harmonics for each of which B and I'_f can be calculated and summed. The products $cU(\Sigma B)(\Sigma I'_f)$ and $c(\Sigma I'_f)^2 \sigma b$ can then be integrated over the generator length

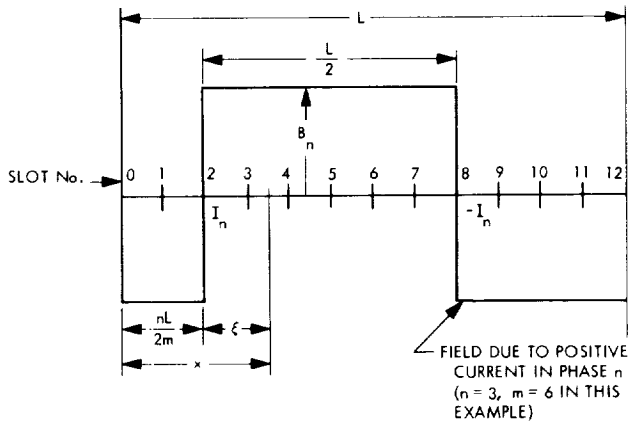


Fig. 5. Winding geometry and field due to one phase

to obtain the input power, ohmic loss, and efficiency with the actual nonsinusoidal waveform.

Figure 5 shows the winding for an m-phase, $2m$-slot generator with $m = 6$. The first phase has coil sides in slots 0 and 6 (or 6 and 12), the second phase in slots 1 and 7, and the sixth phase in slots 5 and 11. To represent the ideal current sheet, the coils carry current $I_n = I_m \cos [(\pi n/m) - \omega t]$, where I_m is the amplitude of the current in each slot and n is the number of the slot. The field produced by one coil is a square wave of magnitude $B_n = \mu_0 i_n/2g$, which can be represented by the series

$$B_n = \frac{2\mu_0 I_n}{\pi g} \sum_{k=1}^{\infty} \frac{1}{k} \sin \frac{2\pi k \xi}{L} \quad (2)$$

for odd k's where k is the order of the harmonic and ξ is the distance from the zero crossing. Substituting for I_n, noting that $\xi = x - (nL/2m)$, and using some trigonometric identities, the empty-channel field due to phase n is

$$B_n = \frac{\mu_0 I_m}{\pi g} \sum_k \frac{1}{k} \left[\sin\left(\frac{2\pi k x}{L} - \omega t - (k-1)\frac{\pi n}{m}\right) + \sin\left(\frac{2\pi k x}{L} + \omega t - (k+1)\frac{\pi n}{m}\right)\right] \quad (3)$$

This equation still describes only the square-wave field of a single coil, but now the field is expressed as a sum of forward-moving and backward-moving harmonics. The total empty-channel field is the summation of Eq. (3) over the m phases. Perform the summation, and the result is

$$B_0(x,t) = \frac{\mu_0 I_m m}{\pi g} \sum_k \frac{1}{k} \sin\left(\frac{2\pi k x}{L} \pm \omega t\right) \quad (4)$$

where the summation is over odd values of k for which either $(k + 1) \bmod (2m) = 0$ or $(k - 1) \bmod (2m) = 0$, the plus sign in Eq. (4) being employed in the former case and the minus sign in the latter. Thus, the empty-channel

magnetic field for six phases, $m = 6$, is

$$B_0(x,t) = \frac{6\mu_0 I_m}{\pi g}\left[\sin\left(\frac{2\pi x}{L} - \omega t\right) + \frac{1}{11}\sin\left(\frac{22\pi x}{L} + \omega t\right) + \frac{1}{13}\left(\frac{26\pi x}{L} - \omega t\right) + \frac{1}{23}\sin\left(\frac{46\pi x}{L} + \omega t\right) + \frac{1}{25}\sin\left(\frac{50\pi x}{L} - \omega t\right) + \cdots\right] \quad (5)$$

Figure 6a shows the empty-channel field at $\omega t = 0$ for six phases with harmonics up to the 49th included ($k = 1$, 11, 13, 23, 25, 35, 37, 47, 49). The wave shape closely approaches the 12 straight-sided steps which can be drawn by simply adding the square wave from each coil, verifying the correctness of Eq. (4).

For the kth harmonic, the wave velocity is $U_{s_k} = \mp\omega L/2\pi k$, the upper and lower signs corresponding to those in Eq. (4), and the product of magnetic Reynolds number and slip is

$$s_k R_{m_k} = \frac{\mu_0\sigma(U - U_{s_k})L}{2\pi k} \quad (6)$$

The magnetic field in the fluid, from Eq. (1), is then

$$B(x,t) = \frac{\mu_0 I_m m}{\pi g}\sum_k \frac{1}{k\left[1 + \left(\frac{s_k R_{m_k} b}{g}\right)^2\right]^{1/2}} \sin\left[\frac{2\pi k x}{L} - \tan^{-1}\left(\frac{s_k R_{m_k} b}{g}\right) \pm \omega t\right] \quad (7)$$

with the k's and the signs selected as described above.

Figure 6b shows the field in the fluid with six phases and harmonics up to the 49th included at operating conditions for the experimental 50-kW generator, namely, $L = 0.1088$ m, $U = 66$ m/s, $s = 0.3$, $\sigma = 1.9 \times 10^6$ mho/m, and $b = g$. The fundamental is shifted downstream 32 deg and reduced in amplitude by 17%, but the harmonics are less affected by the fluid, resulting in a skewing of the steps making up the total field.

The inclusion of harmonics up to the 49th in Fig. 6 provides a close representation of the wave shape ideally produced by a six-phase winding, but in practice the sharp corners are lost due to fringing.

The fluid current produced by each harmonic of the Eq. (7) field is

$$I'_{f_k}(x,t) = \sigma b B_k(x,t)\, U_{s_k} s_k$$

Summing over the harmonics gives the local fluid current I'_f. Numerically integrating $cUBI'_f$ over the generator length gives the input power P_m, and integrating $c(I'_f)^2/\sigma b$ gives the ohmic loss P_r. The efficiency is $\eta_0 = (P_m - P_r)/P_m$. When these calculations are done for $L = 0.1088$ m, $\sigma = 1.9 \times 10^6$ mho/m, $s = 0.3$ (slip of the fundamental), $m = 6$, and $U = 66$ m/s, the efficiency is found to be 0.77 with only the fundamental ($k = 1$) included, 0.71 with the 11th and 13th harmonics, 0.695 with the 23rd and 25th, 0.690 with the 35th and 37th, and 0.687 with the 47th and 49th harmonics included. Thus, a reasonable estimate of the efficiency reduction can be obtained by including only harmonics up to the 25th, and even this may be pessimistic in view of the suppression of harmonics in practice due to fringing.

Figure 7 shows the variation of efficiency with slip at the same conditions as Fig. 6, with only harmonics up to the 25th included. The efficiency is 5 to 10 percentage points below the ideal, sinusoidal value $(1 + s)^{-1}$ down to a slip of 0.1 where the efficiency peaks at 0.75 and then drops to zero. Thus, a larger number of phases than six is required for negligible reduction in efficiency over the ideal continuous winding.

The variation of efficiency with number of phases at a slip of 0.3 (under the same conditions as Fig. 6) is presented in Fig. 8. The efficiency increases from 0.52 with three phases to 0.76 with twelve phases, only one percentage point less than the $(1 + s)^{-1}$ value. Thus, it is concluded that generators with about 24 slots will adequately approach the ideal $(1 + s)^{-1}$ efficiency, and generators with only 12 slots, as in the current experimental generator, will have an efficiency which is reduced by about 5 to 10 percentage points.

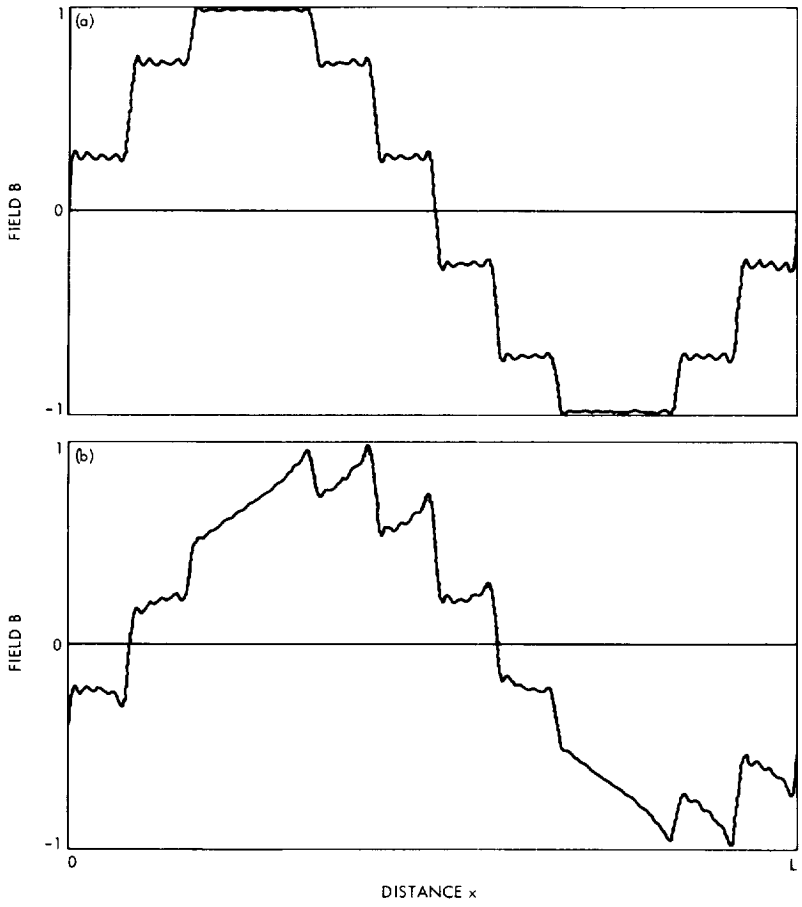


Fig. 6. Theoretical magnetic field in six-phase generator: (a) emply channel, (b) channel with fluid

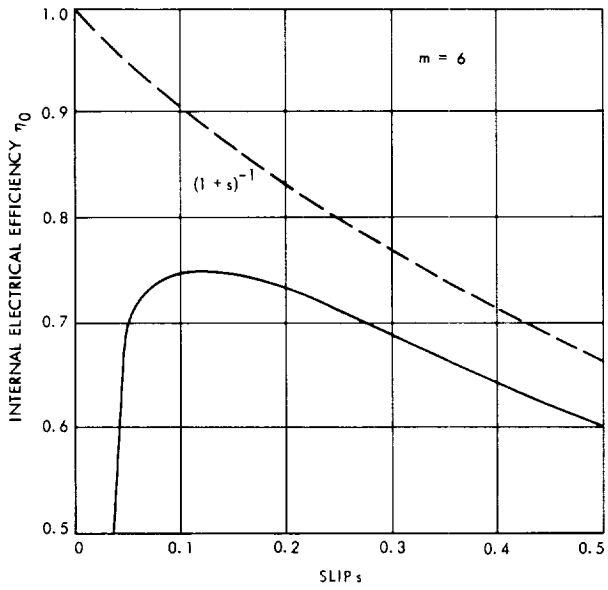


Fig. 7. Theoretical variation of internal electrical efficiency with slip in a six-phase generator with harmonics up to the 25th

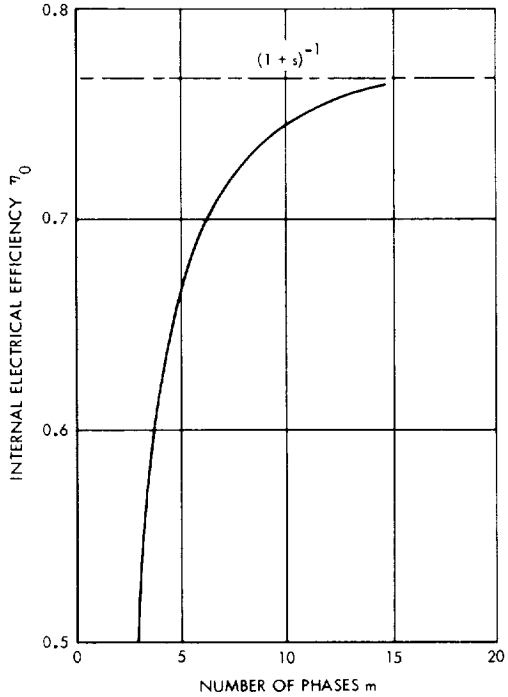


Fig. 8. Effect of number of phases on efficiency

3. High-temperature Bimetal Corrosion Loop

A test loop fabricated of Haynes-Stellite alloy 25 (H-25) completed 100 h of operation in air with flowing lithium at temperatures above 1800°F. Use of this alloy for a 200- to 300-kWe performance test is being investigated because of its low cost relative to refractory metal alloys. All interior surfaces of the loop except the pump duct and weldments were protected from corrosion by Nb-1% Zr tubing or coating. The purpose of this experiment was to evaluate (1) the protective qualities of a coating of vapor-deposited Nb-1%Zr on the interior of the H-25 tubing, (2) the use of Nb-1% Zr tubing to mask the bulk of the H-25 tubing from corrosion, and (3) field welding methods to join sections of the bimetal test loop. Except for the interior coatings and field weldments, the flow system was identical to that described in SPS 37-45, Vol. IV, p. 133.

The maximum velocity during this test was approximately 50 ft/s at the throat of a venturi. The test was interrupted after 82 h of operation above 1800°F by a pump coil failure and again at 102 h by a lithium leak at a location remote from the vapor-deposited test section. Post-test examination revealed no significant evidence of corrosion in the failure region, and the cause of failure is believed to be brittle fracture of the H-25 alloy in a high-stress location. The high-temperature history of the test loop is summarized in Fig. 9. In addition to the 102 h above 1800°F, peak temperatures of 2000°F were reached and a total of 10 h of operation above 1900°F were realized. The total temperature difference ranged from 600 to 800°F.

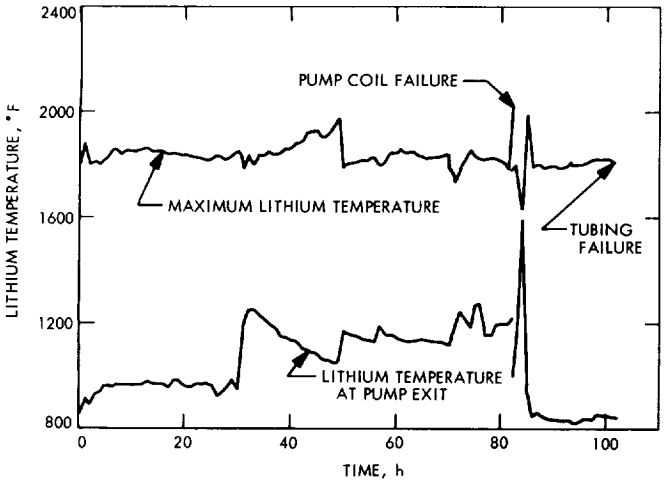


Fig. 9. High-temperature history of bimetal lithium system (H-25;Nb–1 % Zr lining)

Examination of the vapor-deposited section after the test showed the coating to be intact with no spalling or degradation apparent. Complete protection was provided to the Haynes-25 substrate for the period of the test. The surface appearance of the Nb-1%Zr coating had altered considerably during the test.

Figure 10 is a series of scanning electron microscope photographs of the surface at different magnifications. The regular polyhedral surface of the as-deposited coating changed to an amorphous structure with scattered, isolated crystals. The crystals were identified by electron microprobe analysis to be composed of nickel which may have been dissolved from the pumping section and subsequently precipitated from the lithium during the cooling process. The amorphous surface is believed to be the result of a coating or reaction product rather than an attack of the vapor deposit. Electron microprobe analysis identified cobalt, nickel, and chromium to be present on the surface. These elements may have migrated from the pumping section which had no protective coating over the Haynes-25 alloy.

Metallurgical evaluation of sections of the flow system is continuing. However, it appears that inexpensive protection of H-25 alloy from lithium corrosion, by vapor deposition of Nb-1%Zr, is feasible and that present field welding techniques are adequate for at least 100 h of operation at 1800°F.

4. Multistage Cycle Analysis

Analysis of the performance of multistage liquid-metal MHD cycles was continued (SPS 37-51, Vol. III, pp. 120–124). These calculations use the summation of experimental or calculated efficiencies for demonstrated components to arrive at the total cycle efficiency and should be quite realistic. Cycle efficiencies in excess of 10% were realized at 1800°F maximum cycle temperature.

Figure 11a presents the variation of efficiency with condensing pressure for four combinations of the number of stages n, and mass flow ratio of lithium to cesium r_c. The peak efficiency of 10.4% was obtained at a condensing pressure of about 8 psia. This point corresponded to nine stages and an r_c value of 15. The maxima of the curves appear to be very broad with only a ½-percentage point loss in efficiency suffered if the condensing pressure were raised from 8 to 15 psia. If only five stages were used at a mass ratio of 15, an efficiency of 9.6% would be obtained at 6- to 11-psia condensing pressure.

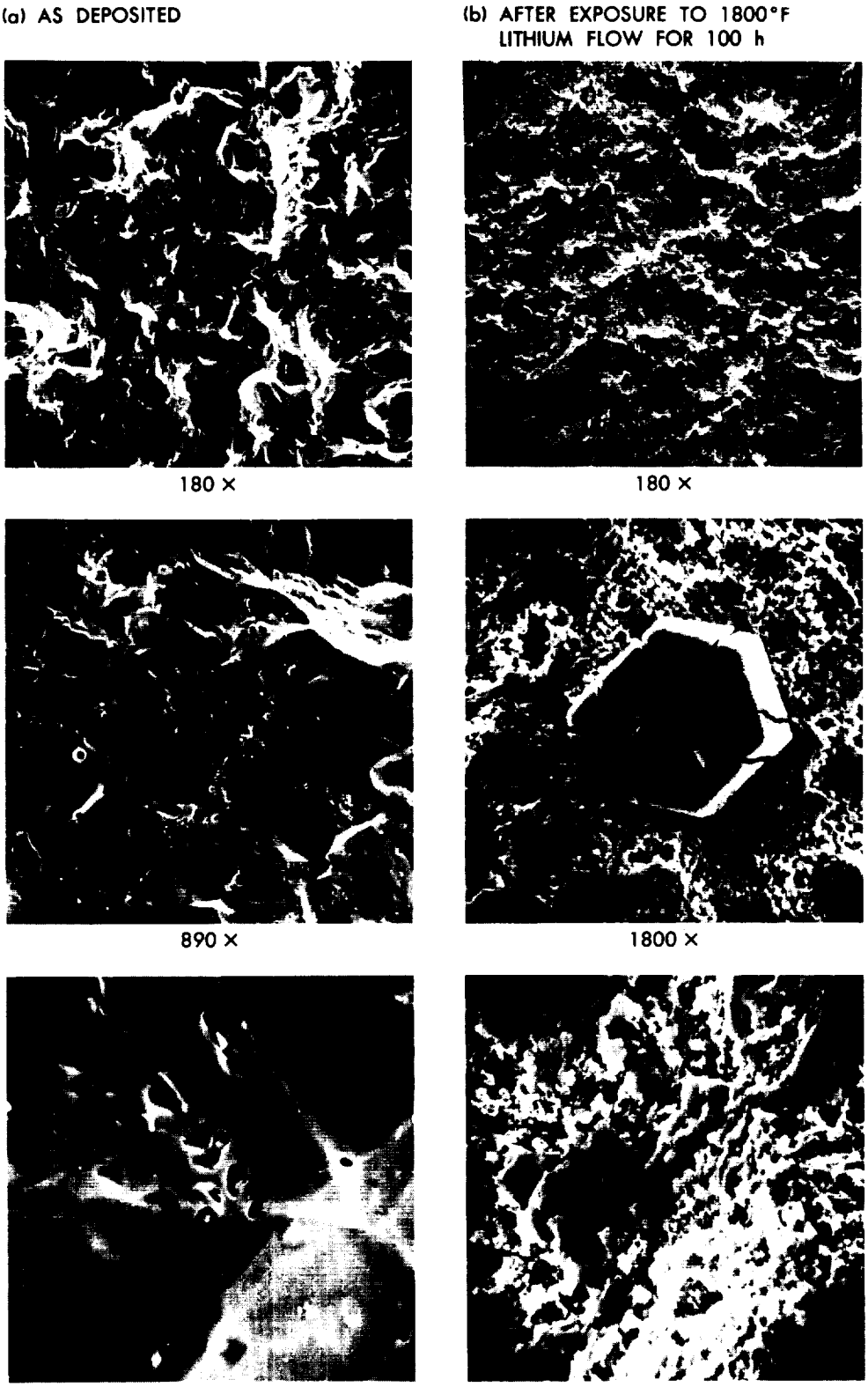


Fig. 10. Surface of vapor-deposited Nb–1% Zr before and after corrosion test

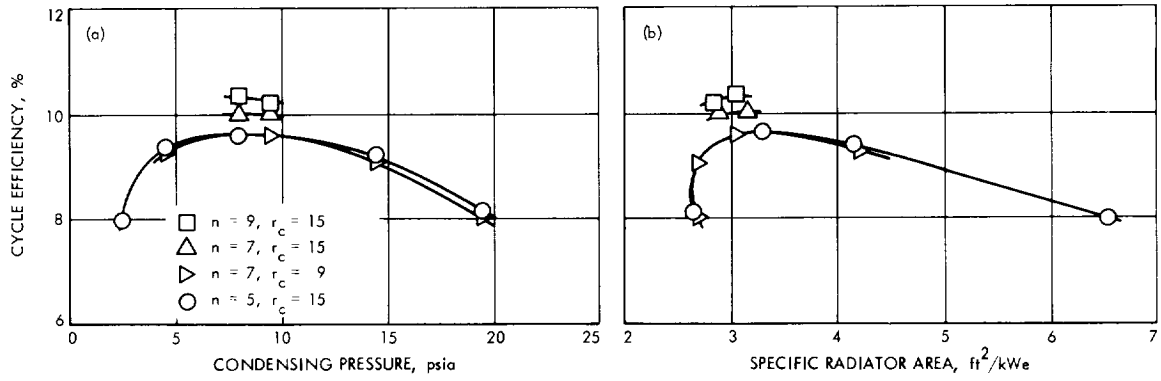


Fig. 11. Multistage liquid–metal MHD cycle efficiency at 1800°F maximum temperature: (a) cycle efficiency vs condensing pressure, (b) cycle efficiency vs specific radiator area

The actual operating point and number of stages chosen will depend upon system tradeoffs of reliability, weight, and configuration to be considered in a future study. As an example of the type of consideration required, Fig. 11b shows the relation of the efficiency to specific radiator area as the condensing pressure is varied. The curve for $n = 5$, $r_c = 15$ shows that the minimum radiator area is not achieved at the maximum efficiency point. The minimum specific area of 2.6 ft²/kWe is achieved at an efficiency of 8.6%. The specific area rises to 3.3 ft²/kWe at the peak efficiency point of 9.6% for this combination of cycle parameters.

An unexpected result of the studies to date is that the maximum cycle efficiency at 2000°F maximum temperature is lower than that for 1800°F. Figure 12 shows a peak efficiency of only 9.4% for seven stages at the optimum value of $r_c = 8$. This can be contrasted to a value of 10.0% for seven stages at 1800°F at the optimum r_c of 15. The lower efficiency is due mainly to the higher partial pressure of lithium vapor and subsequent higher heat load on the condenser. The higher cycle temperature might still be of interest to further reduce the radiator area. For example, the minimum specific area calculated for the 2000°F cycle was 2.1 ft²/kWe compared to the minimum calculated value of 2.6 ft²/kWe for the 1800°F cycle. However, variation of the parameters is incomplete for the lower cycle temperature and further reductions may be possible. Performance calculations are continuing for further variations in the number of stages, mass flow ratio, and power level.

D. Dynamic Gas Effects on the Breakdown Potential of Helium, *J. A. Gardner*

An experimental investigation was conducted to determine the influence of gas velocity and pressure on the

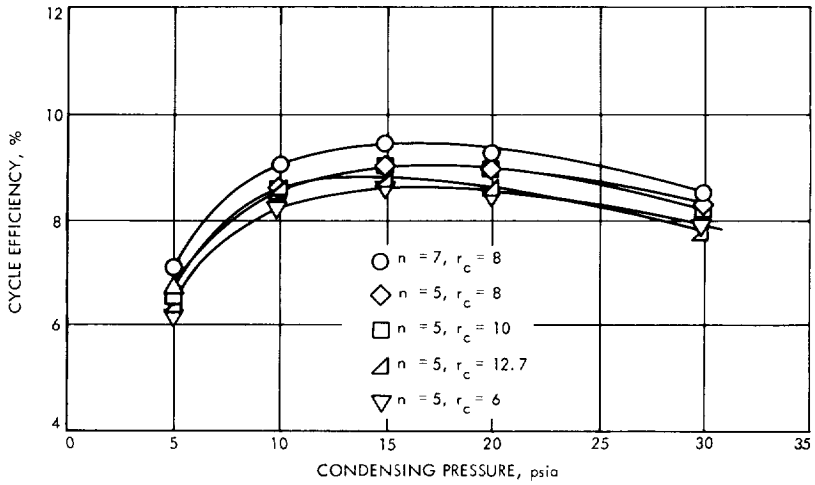


Fig. 12. Multistage liquid–metal MHD cycle efficiency at 2000°F maximum temperature

electrical breakdown potential of helium between parallel flat-plate electrodes. These experiments were performed with helium to determine whether or not the velocity–pressure effects observed with argon might also be exhibited with another monatomic gas. The apparatus was the same as that used to conduct the experiments with argon reported in SPS 37-47, Vol. III, pp. 143–148 and in Ref. 1.

The results of the argon experiments indicated that the breakdown potential was reduced by a factor of 2 to 3 in the pressure range between 50 and 200 torr, exhibiting a significant reduction at velocities of the order of 25 to 100 ft/s. After the initial reduction in breakdown potential at the onset of a transverse gas motion, a further increase in the gas velocity had a lesser additional effect. Druyvestyn and Penning (Ref. 2) reported the breakdown potentials for static conditions of several monatomic and diatomic gases in the form of a Paschen curve of breakdown potential versus the product of pressure and electrode separation distance.

The parallel flat-plate electrodes were made of copper and installed downstream of a lucite nozzle. They were 1 in. square, ½ in. thick, and were separated by a distance of ½ in. The exit of the nozzle was ½ in. high and 2 in. wide. Helium entered the space between the electrodes at ambient temperature.

The high-voltage dc power supply used to provide the breakdown potential to the electrodes consisted of a primary controlled transformer, a fullwave rectifier, and two 4-μf capacitors for filtering.

The procedure for conducting these breakdown experiments consisted of, first, setting the desired gas flow conditions. Then, the primary voltage of the high-voltage power supply was increased until the onset of breakdown was observed on an oscilloscope and recorded on polaroid film. The output of the power supply was reduced for measurement by the use of a 10,000/1 precision voltage divider connected between the power supply and the inputs of the oscilloscope.

Breakdown potentials for helium are shown in Fig. 13. Results were obtained over gas velocity ranges of 0 to 525 ft/s at 50 torr, 0 to 425 ft/s at 100 torr, 0 to 160 ft/s at 200 torr, and 0 to 75 ft/s at 400 torr. The gas velocities

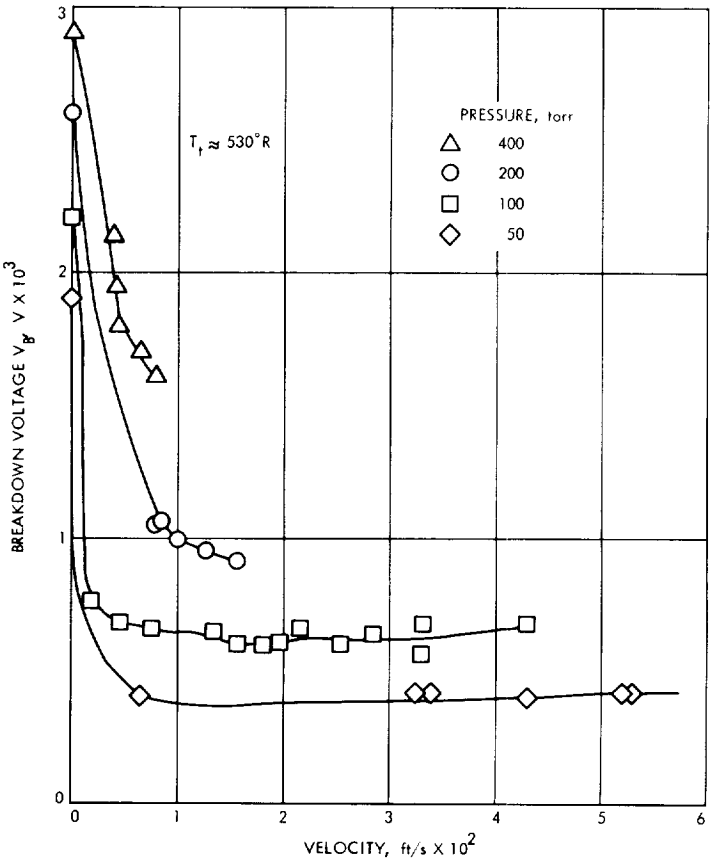


Fig. 13. Influence of velocity and pressure on breakdown voltage for helium flow between flat parallel-plate electrodes

between the electrodes were computed from measurements of the total mass flow rate, the pressure and temperature of the gas, and the cross-sectional area of the throat of the nozzle just upstream of the parallel flat-plate electrodes. Uniformity of the velocity at the nozzle exit, in the absence of a discharge, had been determined by pitot tube traverses prior to the argon flow experiments.

As shown in Fig. 13, the breakdown potential for helium at all pressures tested is significantly below that for a stagnant gas, even for comparatively low gas velocities. Furthermore, once this reduction occurred, an increase in the gas velocity had a smaller effect. In the 50- and 100-torr pressure regime where data were taken over a larger velocity range, leveling off appears to develop at approximately 420 and 640 V for the 50- and 100-torr conditions, respectively.

A comparison of the breakdown potentials between helium and argon at a pressure of 100 torr is shown in Fig. 14. The effect of velocity exhibits the same trends for the two gases with the breakdown potential for argon being higher at all velocities.

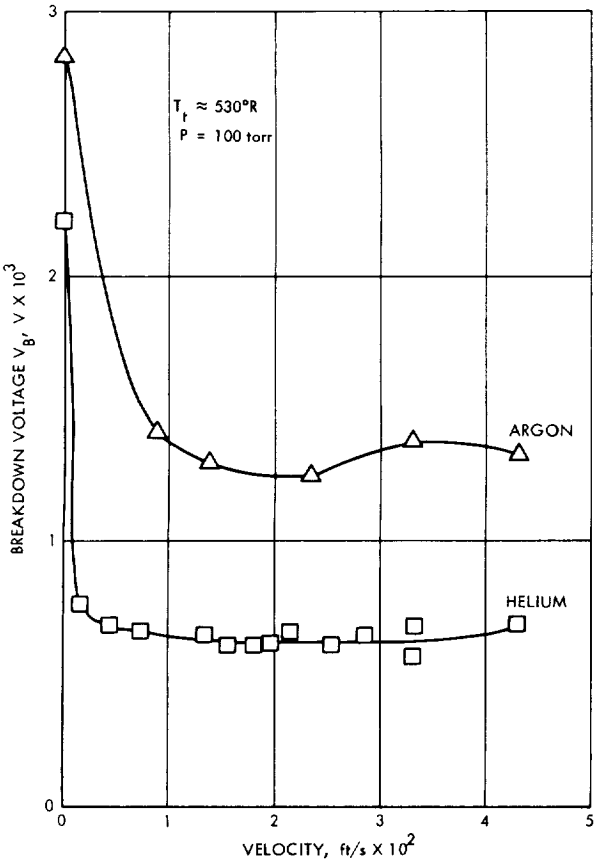


Fig. 14. Comparison between helium and argon of the velocity effect on breakdown potential at a pressure of 100 torr

References

1. Gardner, J. A., "Effects of a Dynamic Gas on Breakdown Potential," *AIAA J.*, Vol. 6, No. 7, pp. 1414–1415, July 1968.
2. Druyvestyn, M. S., and Penning, F. M., "The Mechanism of Electrical Discharges in Gases of Low Pressure," Rev. Mod. Phys., Vol. 12, p. 87, 1940.

E. Lithium–Boiling Potassium Test Loop Runs With Reactor Simulator, *H. Gronroos and G. Kikin*

1. Introduction

Phase III operations of the lithium–boiling potassium Rankine-cycle test loop at JPL have been completed and represent the final experiments that were performed before the dismantling of the experimental facility for metallurgical examinations. Earlier phases were devoted to thermal-hydraulic and stability experiments utilizing the co-annular boiler. The final phase of operations was performed with a shell side boiling, cross-flow boiler and with the analog nuclear-reactor-simulator coupled to the test loop.

The primary purpose of the test loop is to investigate overall transient and steady-state characteristics of a two-loop Rankine-cycle pilot plant, which contains all the essential components of a Rankine space power plant and approximately simulates velocities, temperatures, pressures, transit times, and heat fluxes in the range of actual system interest. Transient investigations include startup, load perturbations, shutdown, and the effects of various control concepts on system performance. The primary heat source is direct-current-resistance heat generation in a section of tube wall and liquid metal. This method of heat generation lends itself to nuclear reactor simulation using an analog computer for observation of power and temperature responses to various system perturbations.

This article briefly discusses the formulation of the analog representation of the nuclear reactor heat source, its coupling to the lithium–boiling potassium test loop, and some of the results obtained. Later JPL publications will cover additional phases of test loop operations and provide further detail and results.

2. Test Loop

The test loop has been described in SPS 37-44, Vol. IV, pp. 150–163, and Refs. 1 and 2. An isometric view of the

test loop is shown in Fig. 15; Table 4 summarizes the essential parameter data. A direct-current resistance-heated helical coil provides the heat input to the lithium primary loop fluid. The resistance coil is driven by a power supply which is in turn controlled by a Robertshaw controller. Potassium vapor generated in the boiler drives a turboalternator, is condensed in a radiating condenser, and then returns via a preheater to the boiler inlet. The design of the loop components is detailed in Ref. 1; however, a shell side boiling, cross-flow boiler was substituted for the original co-annular boiler described in Ref. 1. This design modification essentially eliminated the boiling instabilities which had been experienced with the earlier boiler design.

Obviously, the dynamic characteristics of the lithium heater coil with power supply and associated controller do not closely correspond to those existing in a nuclear reactor. Therefore, not much is gained by directly driving the controller with a signal obtained from simulation of the neutron kinetics equations. However, a reference reactor model may be chosen such that the temperature drop between the reactor coolant outlet and inlet temperatures is the same as that across the heater coil. One

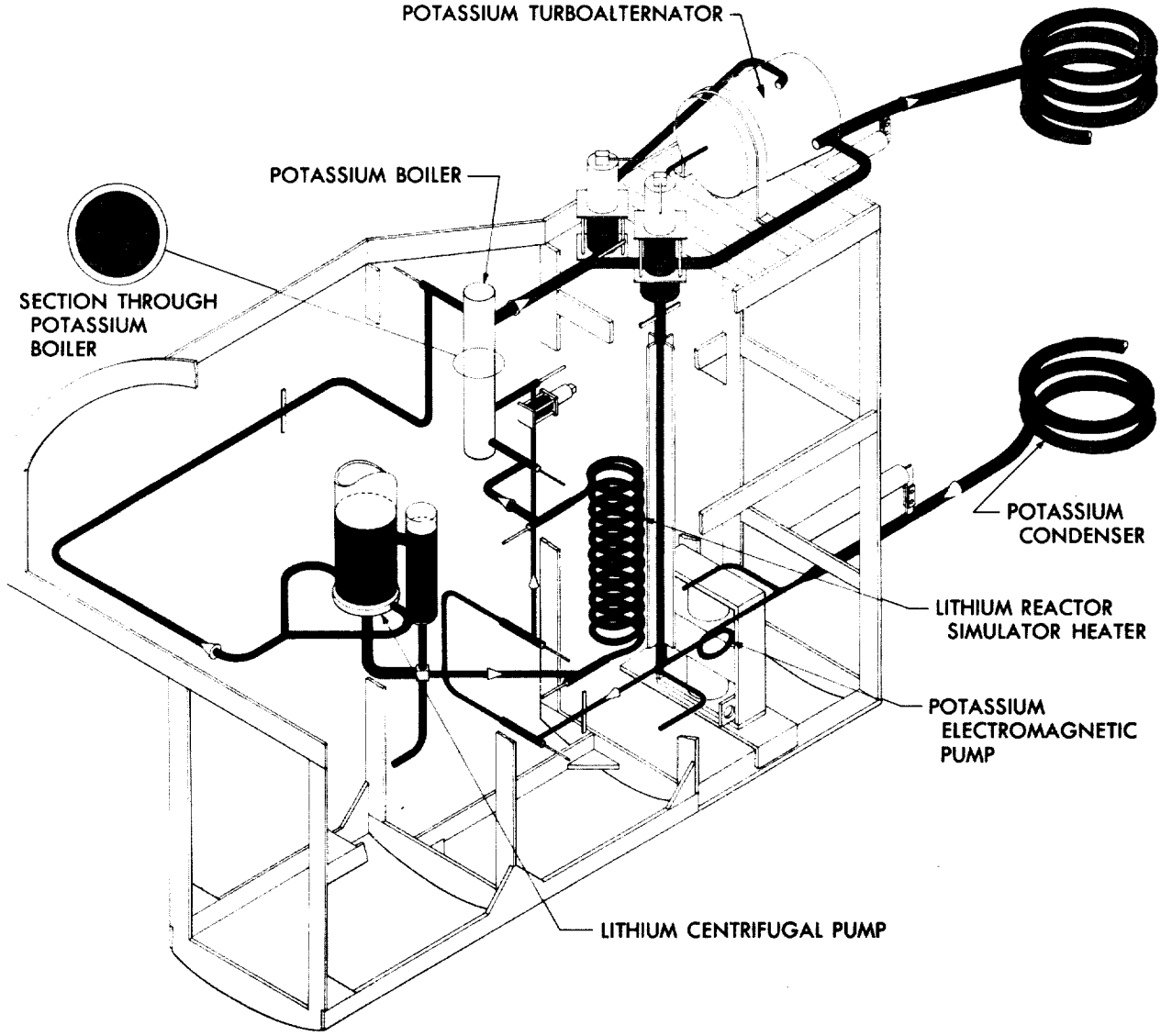


Fig. 15. Isometric of lithium–boiling potassium test loop

Table 4. 30 kW–2100°F loop operating conditions

Item	Operating conditions	Supplier and type
Lithium (liquid)		Foote Mineral Company
Flow rate	5 gpm	
Temperature	2100°F	
Pressure	Up to 20 psig	
Potassium		MSA Research
Flow rate	0 to 1 gpm	
Temperature	1500 to 2000°F	
Pressure	Up to 200 psig	
Centrifugal pump (lithium)	Up to 10 gpm at 100 ft 2100°F service	Byron–Jackson
Em pump, dc (potassium)		MSA Research Cb–1Zr
Temperature	1500°F	
Head-flow, nominal	90 psi–150 lb/h	
Swing gate valves (bellows seal)		Valcor Engineering
Potassium	2000°F boiling potassium service	3/8 in.
EM flowmeters		MSA flowmeter FM-4
Lithium	2100°F service	5/8 in. OD 1/16 in. wall
Potassium	1500°F service	Cb–1Zr duct 3/8 in. OD Cb–1Zr duct
Potassium boiler	Loop design conditions	JPL
Turboalternator		
Inlet temperature	1900°F	Aeronutronic
Exhaust temperature	1500°F	Cb–1Zr case
Weight flowrate	0.0324 lb/s	Molybdenum turbine wheel, shaft, alternator rotor
Design speed	12,000 rpm	
Electrical output	1.0 kW	SS 316 alternator housing
Dump tanks and valves (Argon, vacuum, and fill)	500°F service	Material SS 304 and 321

can then attempt to design an on-line control element that gives the correct reactor transfer function for the experimental assembly. In other words, if $T_o(s)/n(s)$ is the transfer function between outlet temperature and neutron density in the reference reactor, a control element is determined so that the transfer function between the signal to the power supply controller and coil outlet temperature, $T_{co}(s)/u(s)$, is equal to $T_o(s)/n(s)$. The control element is most conveniently programmed on an analog computer on which the reference reactor is also simulated. This makes it possible to define the reactor state variables and the various reactivity coefficients. The common state variables between the test loop and reference reactor are the measured heater coil coolant inlet and outlet temperatures.

In attempting to implement the above scheme, a difficulty arises from the need to realize the transfer function between the reactor outlet and inlet temperatures, $T_o(s)/T_i(s)$, and not only the transfer function $T_o(s)/n(s)$ mentioned above. Also, the thermal power level is only about a twentieth of that in the smallest practical space nuclear powerplant. Scaling up proportionately from the power level used in the experimental assembly implies that the full-scale plant would consist of the proportionate number of parallel loops, while in a true system probably no more than two parallel loops would be used. In addition, since the loop components were not manufactured to space application specifications, the time constants only approximate those to be expected in a real space powerplant.

Despite the previously mentioned fundamental constraints, the inclusion of a reactor simulator in the test loop significantly aided the investigations of the dynamic behavior of nuclear liquid-metal Rankine-cycle powerplants. A gross view of the kinetics was obtained, and by comparisons to other similar analytical and experimental studies, identification of the significant parameters may be inferred.

3. Analog Simulation

In order to determine the necessary compensating element for the realization of a desired transfer function for the power supply–controller assembly, the transfer function of the power supply with controller, $K_{RT}G_{RT}$, was measured separately. The following was obtained for the open-loop system:

$$K_{RT}G_{RT} = \frac{3060}{s(s^2 + 10.2s + 473)} \qquad (1)$$

To this is added the equations describing the thermal balance in the coil. This set is compared with the set obtained from describing the reference reactor system.

A variety of techniques is available to design the compensation such that the desired transfer function defined by the reference reactor is obtained for the experimental assembly. The uncertainties present do not warrant any extended effort in this regard; and, therefore, linear state variable feedback techniques were used. The necessary

state variables are fed back through constant-gain frequency-independent elements so that the resulting overall transfer function is the desired one (Refs. 3 and 4). If some important state variable is not measurable, it can be generated through its known relationship to a measurable variable.

The desired transfer function between reactor outlet temperature and neutron density was calculated to be

$$\frac{T_o(s)}{\delta n(s)} = \frac{3944}{(s+8.86)(s+98.8)(s+0.8)(s+0.2761)} \quad (2)$$

The prompt jump approximation was used with the point kinetics equations. The thermal balance equations were spatially lumped so that a low-order model was obtained. Equation (2) was determined by assuming constant reactor inlet temperature and, consequently, constant heater coil inlet temperature.

Only the coolant temperatures and power supply voltage were readily measurable. This was accomplished using amplifiers with common mode rejection and filters. These signals were fed to a Systron-Donner model 20 desk analog computer on which the neutron kinetics and

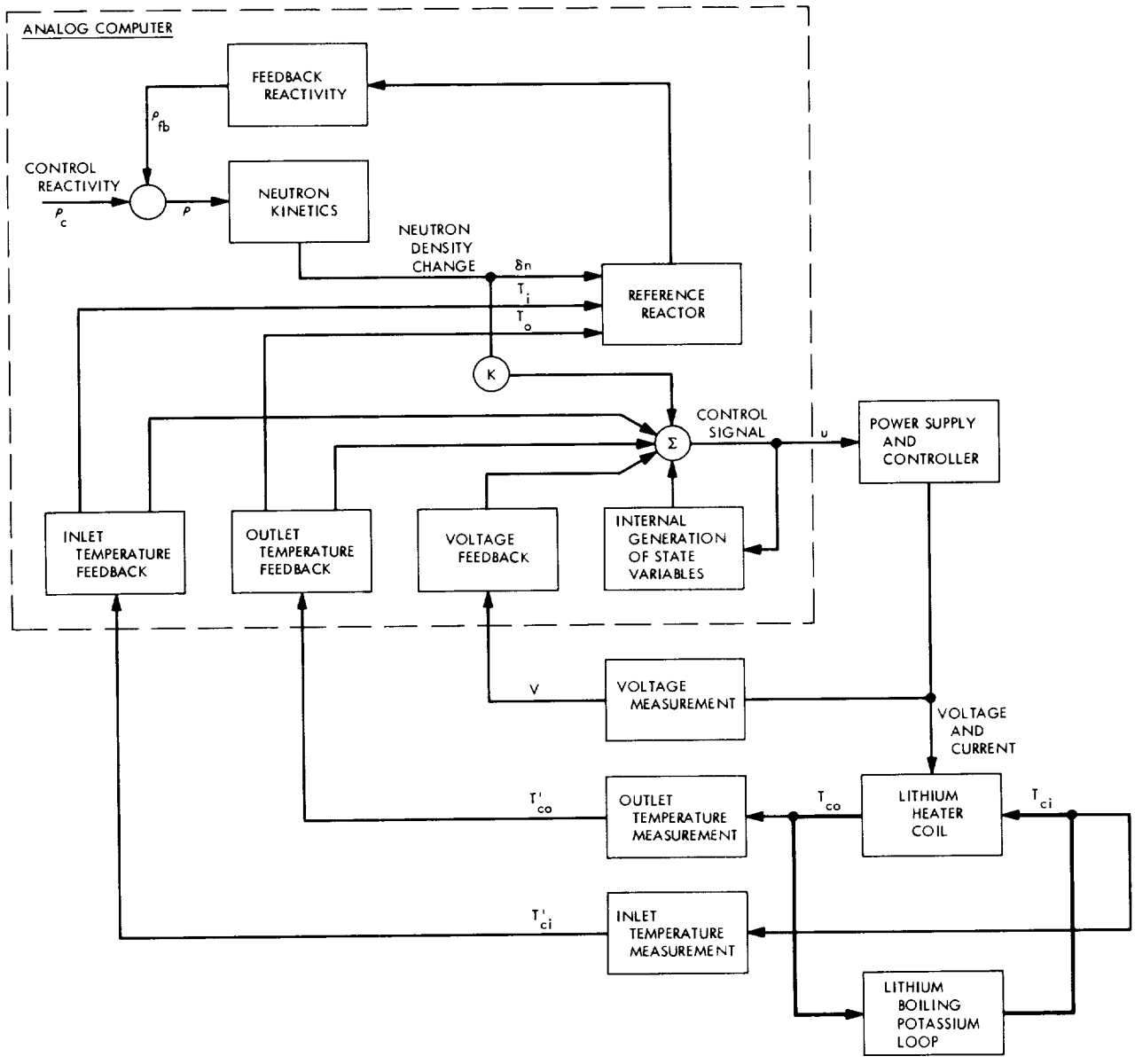


Fig. 16. Schematic of reactor simulator

reactor thermal-balance equations were programmed. Equations defining additional non-measurable state variables were also programmed on this computer. Because of limitations of available equipment, compromises had to be made in setting up the simulation by deleting the least important states. The temperature coefficients of reactivity, which make up the internal feedback loops in the reactor core, were not calculated; but their range of values was estimated and reasonable values were set on the coefficient potentiometers.

Figure 16 illustrates schematically the obtained simulation setup. The analog computer was brought on line by balancing the control signal to zero and then activating a switch on the main control board, which also contained the safety logic. By setting a gain adjustment potentiometer, the same percentage variation in neutron density as that observed for the heater coil power supply was obtained.

4. Results

A variety of cases was investigated with the simulator on line, including non-boiling loop conditions, boiling start-up, reactivity perturbations, and perturbations in the loop components. Figure 17 shows the response to a step change in reactivity. The long time constant for the loop is evident, while the reactor simulator responds quickly. It was observed that small perturbations in the secondary side of the thermal loop had negligible effects on the state of the reactor simulator. Only large changes in potassium pump power produced appreciable variation, as shown in Fig. 18. Due to an electrical winding failure within the turboalternator thermal power demand, variations could not be readily introduced.

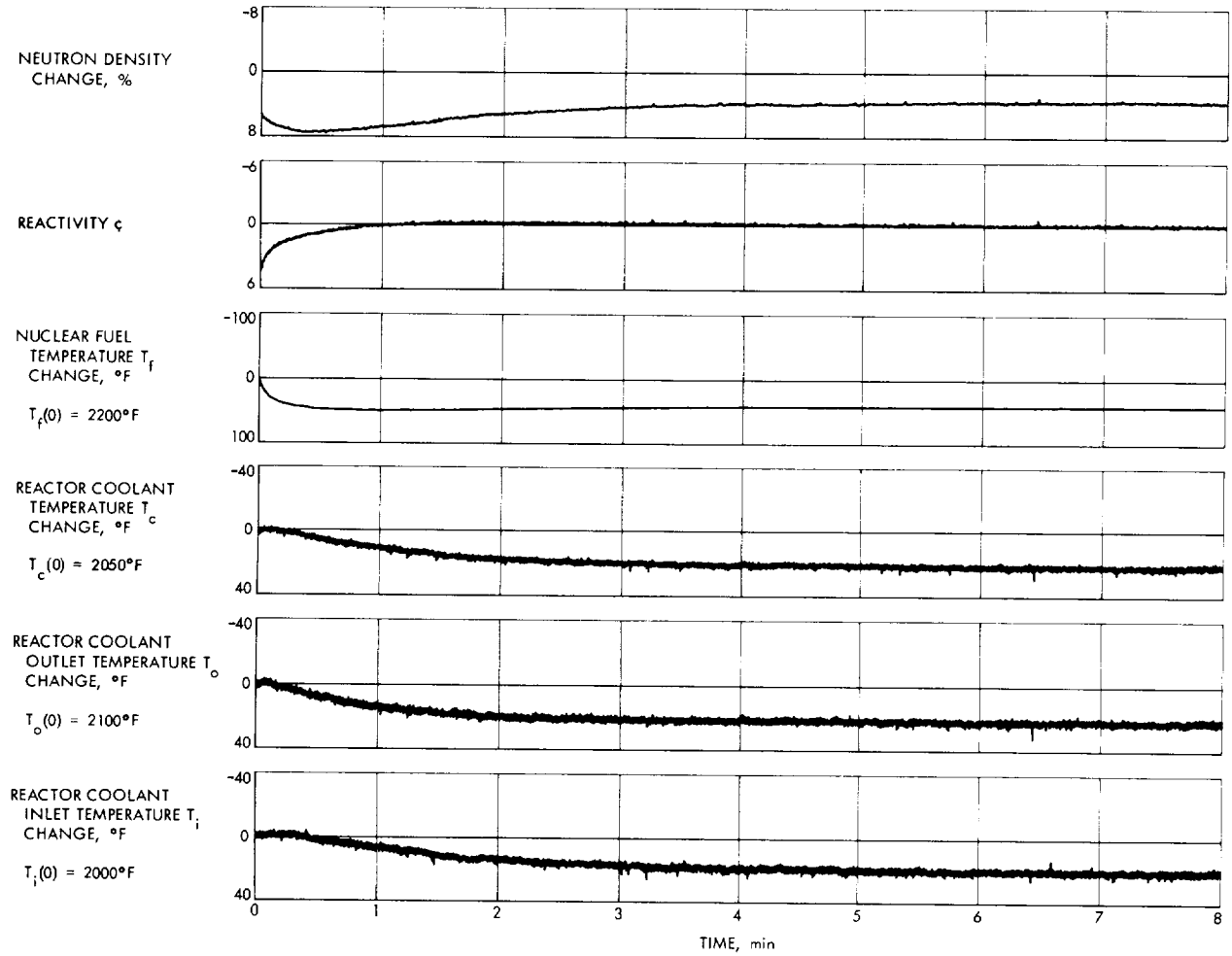


Fig. 17. Response of reactor simulator to a +5¢ step reactivity perturbation

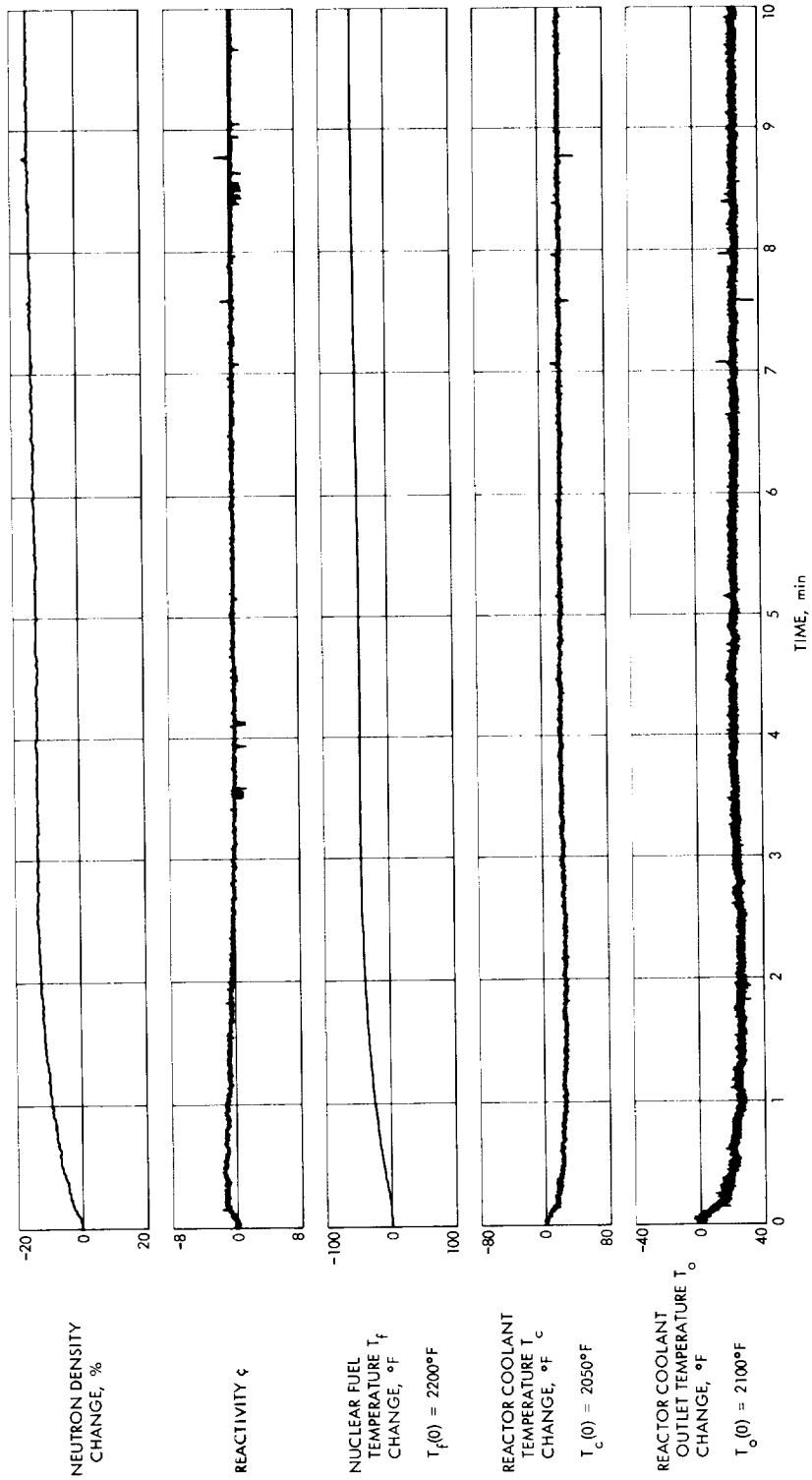


Fig. 18. Response of reactor simulator to a step perturbation in potassium pump current from 2000 to 1300 A

Comparison of the responses with those obtained from other studies on corresponding nuclear systems shows that the results are similar. No unusual behavior was encountered in the dynamic behavior when the reactor simulator was on line, and the loop operated in a very stable manner. The only oscillatory behavior experienced occurred at significantly off-design conditions where the boiler saturation temperature and pressure were greatly below the nominal values. At these off-design conditions, unstable oscillatory boiling is more prone to occur due to liquid-metal superheat and boiling nucleation problems. These boiler exit-vapor temperature oscillations were only observed in the secondary loop; no oscillations were observed in the primary loop.

References

1. Kikin, G. M., et al., *Lithium–Boiling Potassium Test Loop, Interim Report*, Technical Report 32-1083. Jet Propulsion Laboratory, Pasadena, Calif., Sept. 15, 1966.
2. Davis, J. P., and Kikin, G. M., "Lithium–Boiling Potassium Rankine Cycle Test Loop Operating Experience," paper presented at the 1967 Intersociety Energy Conversion Engineering Conference, August 13–17, 1967, Miami Beach, Fla.
3. Weaver, L. E., and Vanesse, R. E., "State Variable Feedback Control of Multiregion Reactors," *Nucl. Sci. Eng.*, Vol. 29, pp. 264–271, 1967.
4. Herring, J. W., et al., "Design of Linear and Nonlinear Control Systems Via State Variable Feedback with Application in Nuclear Reactor Control," Engineering Experiment Station Report. University of Arizona, Tucson, Ariz., Feb. 1967.

XII. Liquid Propulsion

PROPULSION DIVISION

A. The Reaction of OF_2 With B_2H_6: Rate of Formation of BF_3 and Consumption Rates of OF_2 and B_2H_6, *R. A. Rhein*

A preliminary experimental study of the OF_2/B_2H_6 reaction (SPS 37-42, Vol. IV, pp. 70–80, 103–104) indicated that at 20-torr partial pressure of each reactant and at ambient temperatures, the main products were BF_3, H_2, and solid materials; the initial rate of formation of BF_3, at 300°K, was found to vary with the initial concentration of the reactants as follows:

$$\left[\frac{d}{dt}(P_{BF_3})\right]_0 = 5.2 \times 10^{-4} (P_{OF_2})_0^{2.15} (P_{B_2H_6})_0^{-0.43} \quad (1)$$

where concentration is in units of torr partial pressure (at 300°K), and time is in minutes.

The experimental data from which the coefficients in Eq. (1) were derived have since been reevaluated, using a computer program designed to process spectrophotometric data. These results of recently measured (computer-processed) coefficients for consumption rate equations for the reaction of OF_2 with B_2H_6 are presented in this report.

The experimental apparatus and procedures used for the consumption rate measurements were the same as for the previous experiments and are described in SPS 37-46, Vol. IV, pp. 173–180. In addition to those procedures, however, a heated cell was used for measurements at elevated temperatures, ranging in 6-deg intervals from 300 to 330°K. The spectrophotometric transmittance measurements were made at the 6.88-μm band of BF_3, the 12.1-μm band for OF_2, and the 6.18-μm band for B_2H_6.

The experimentally derived values for the rates of B_2H_6 and of OF_2 consumption and of BF_3 formation are plotted versus initial reactant concentration or temperature in Figs. 1 to 5. When these data are used as input for the computer program, the coefficients for the rate expressions are determined, and the several rate equations become

$$\left[\frac{d}{dt}(P_{BF_3})\right]_0 = 2.16 \times 10^{-3} (P_{OF_2})_0^{1.699} (P_{B_2H_6})^{-0.432} \quad (2)$$

$$-\left[\frac{d}{dt}(P_{B_2H_6})\right]_0 = 2.82 \times 10^{6} (P_{OF_2})_0^{2.200} (P_{B_2H_6})^{-0.562} \exp(-11{,}466.9/RT) \quad (3)$$

$$-\left[\frac{d}{dt}(P_{OF_2})\right]_0 = 0.294 (P_{OF_2})_0^{1.624} (P_{B_2H_6})^{-0.024} \exp(-2822.5/RT) \quad (4)$$

where P is concentration in units of torr at 300°K; t is time, min; T is temperature, °K; and R is 1.987 cal/mole-°K.

It is significant that rate magnitudes of B_2H_6 consumption and BF_3 formation are dependent inversely on B_2H_6 concentration, while the OF_2 consumption rate is independent of B_2H_6 concentration. This rate dependence, or lack of it, must be accounted for by any mechanism proposed for the OF_2/B_2H_6 reaction.

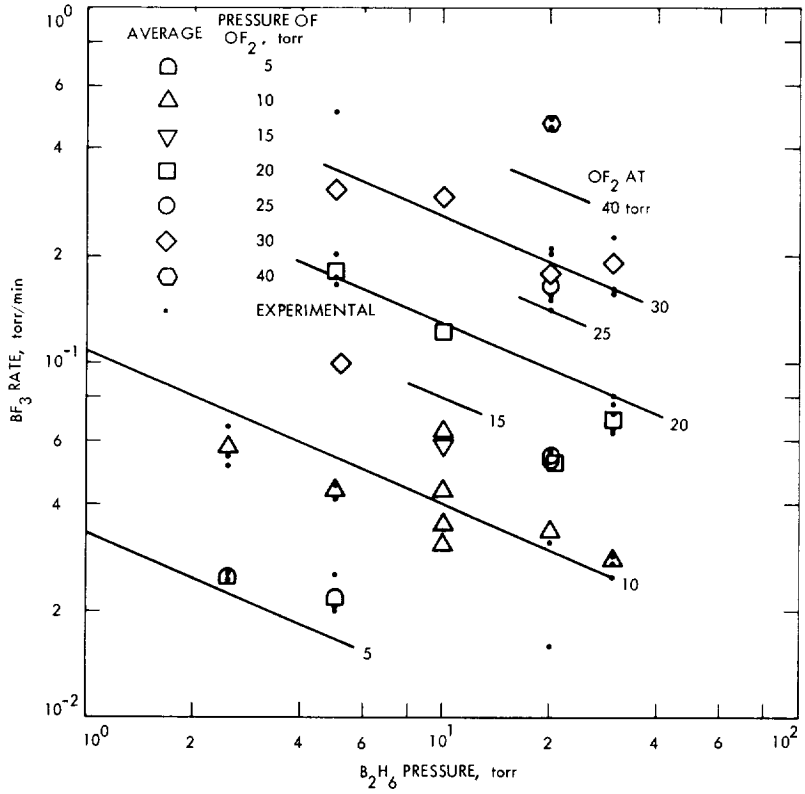


Fig. 1. Isothermal (300°K) rate of formation of BF_3 versus B_2H_6 pressure from the reaction of B_2H_6 with OF_2

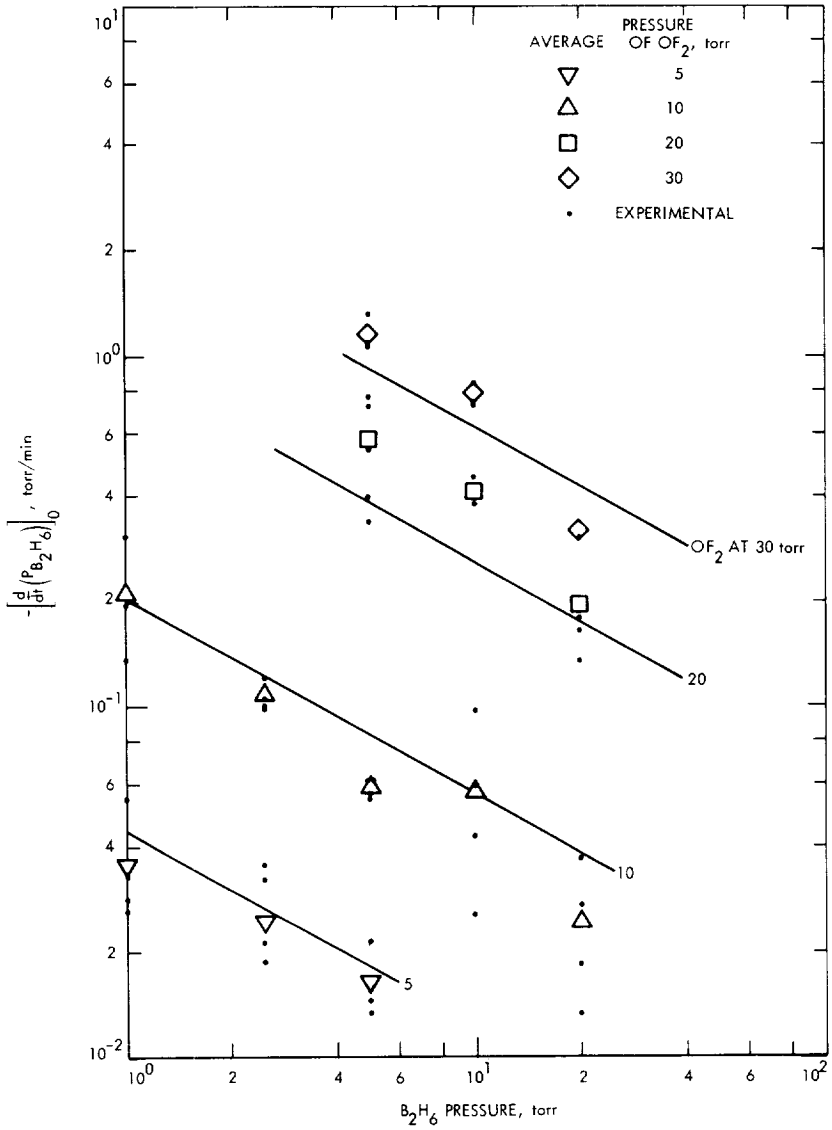


Fig. 2. Isothermal (300°K) consumption rate of B_2H_6 as a function of B_2H_6 concentration

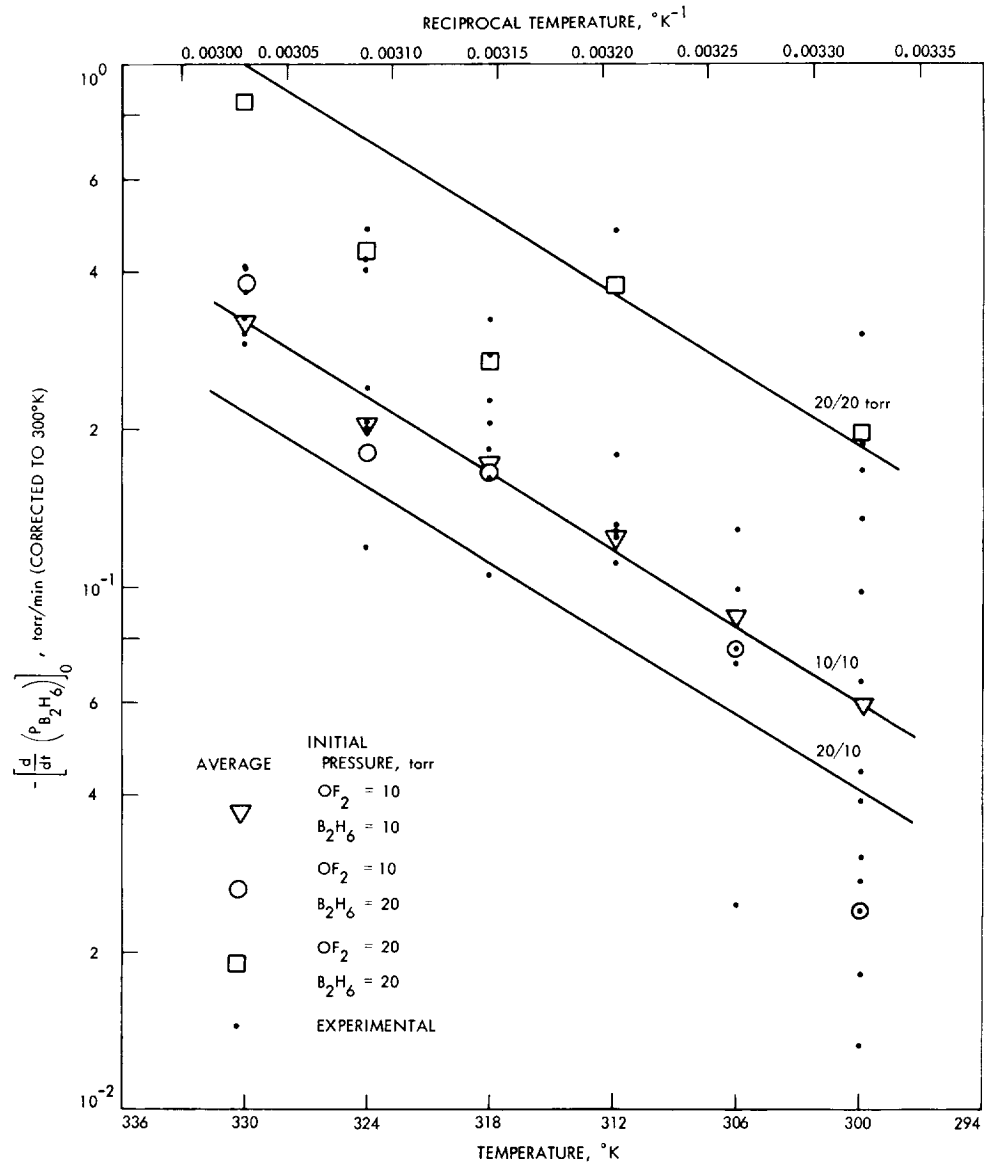


Fig. 3. Consumption rate of B_2H_6 as a function of temperature

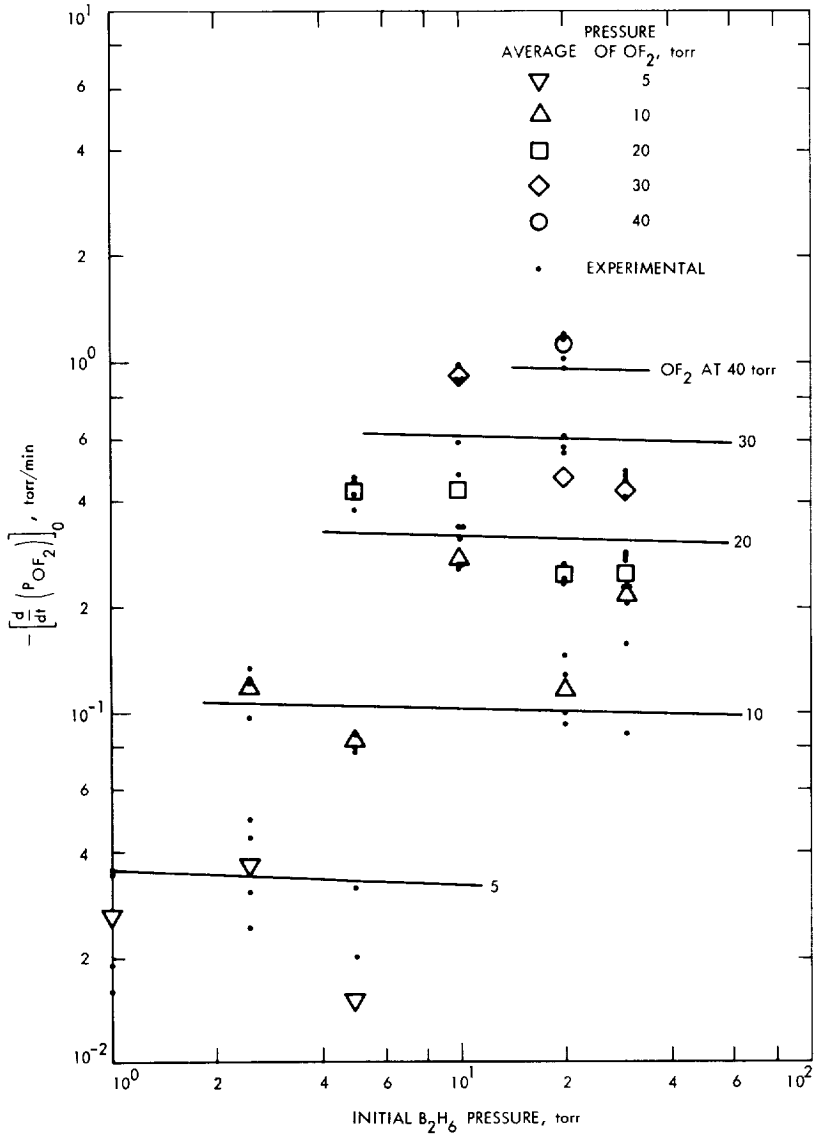


Fig. 4. Isothermal (300°K) consumption rate of OF_2 as a function of OF_2 concentration

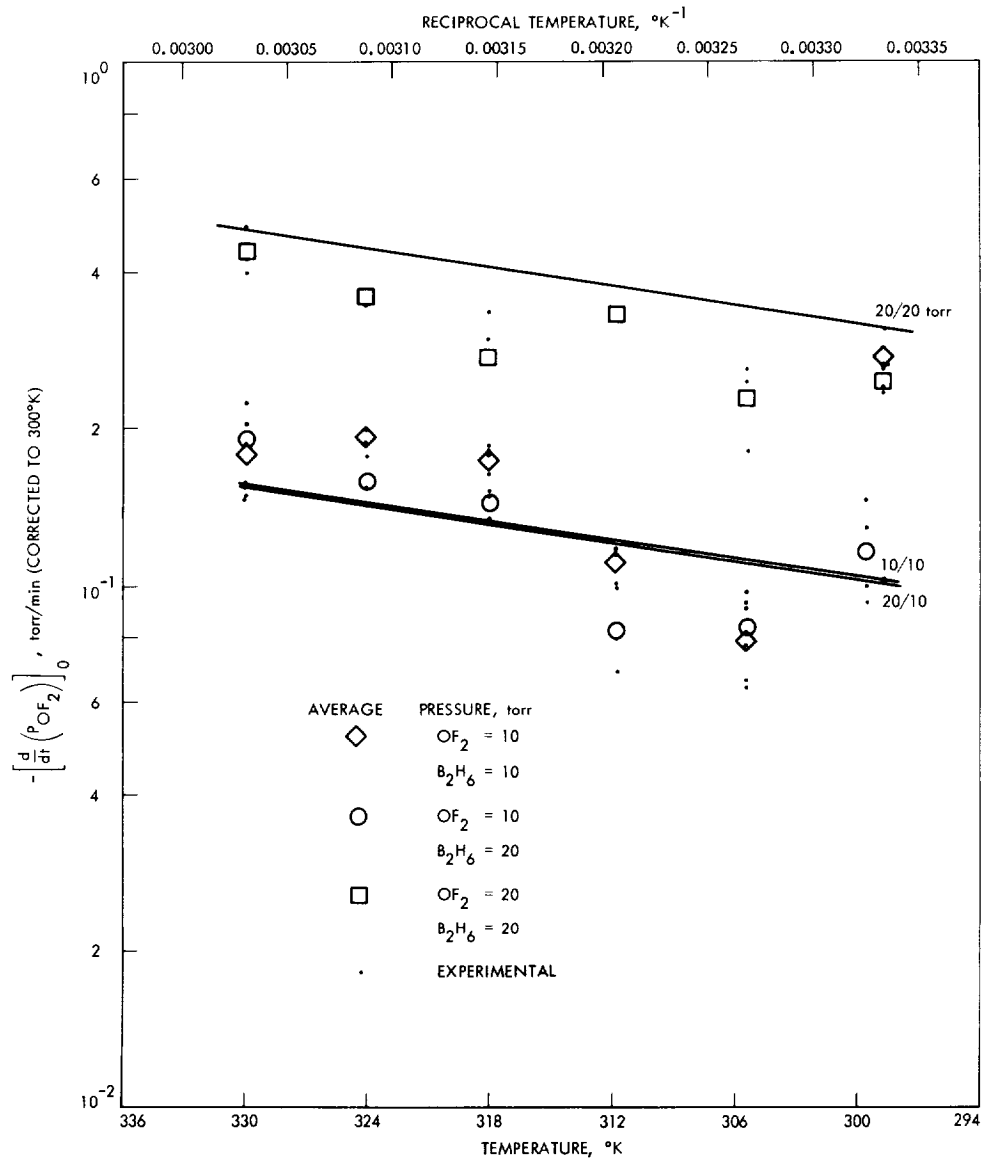


Fig. 5. Consumption rate of OF_2 as a function of temperature

XIII. Space Instruments

SPACE SCIENCES DIVISION

A. High- and Low-Field Operation of the Helium Vector Magnetometer, *F. E. Vescelus*

The phenomenon of optical pumping in metastable helium underlies the operation of the helium vector magnetometer (Ref. 1). The transmission of 1.083-μm resonant radiation through a helium plasma is dependent on the angle between the magnetic field in the plasma and a circularly polarized, resonant, collimated light beam. The resultant light through the cell can then be monitored by a light detector sensitive to the 1.083-μm region, yielding a signal S from the detector of

$$S = k \cos^2 \theta \tag{1}$$

where θ is the angle between the light beam and the magnetic field, and k is a constant dependent on parameters of the measuring system. The sensor components are surrounded by a triaxial set of Helmholtz coils which are used to generate a rotating magnetic vector and to null out the ambient magnetic field at the sensor.

As shown in SPS 37-30, Vol. IV, pp. 131–134, the signal at the detector, using Eq. (1) and vector considerations, is

$$S = k \left[\frac{\left(\cos \omega t + \frac{H}{H_s} \cos \phi \right)^2}{\left(\cos \omega t + \frac{H}{H_s} \cos \phi \right)^2 + \left(\sin \omega t + \frac{H}{H_s} \sin \phi \right)^2} \right] \tag{2}$$

where

H = magnitude of ambient field

H_s = magnitude of sweep vector

ϕ = angle between light axis and H

ωt = angle between light axis and H_s

When the sensor is used as a null device, with the feedback coils reducing the ambient field to a residual "error" field H, then $H \ll H_s$, and Eq. (2) can be simplified by discarding $(H/H_s)^2$ terms and approximating $1/(1 + a)$ by $1 - a$, yielding

$$S = \frac{k}{2} \left[1 + \cos 2\omega t + \frac{H}{H_s} \cos \phi \, (\cos \omega t - \cos 3\omega t) - \frac{H}{H_s} \sin \phi \, (\sin \omega t + \sin 3\omega t) \right] \tag{3}$$

Equation (3) indicates that the components of the ambient field cause perturbations synchronous with their corresponding sweep vectors. Also, since H/H_s is small compared to unity, the major ac portion of the signal is the second harmonic of the sweep frequency. This signal must therefore be rejected and the signal at the sweep frequency processed.

Figure 1 shows the helium vector magnetometer block diagram. The rotating magnetic vector sweeps about

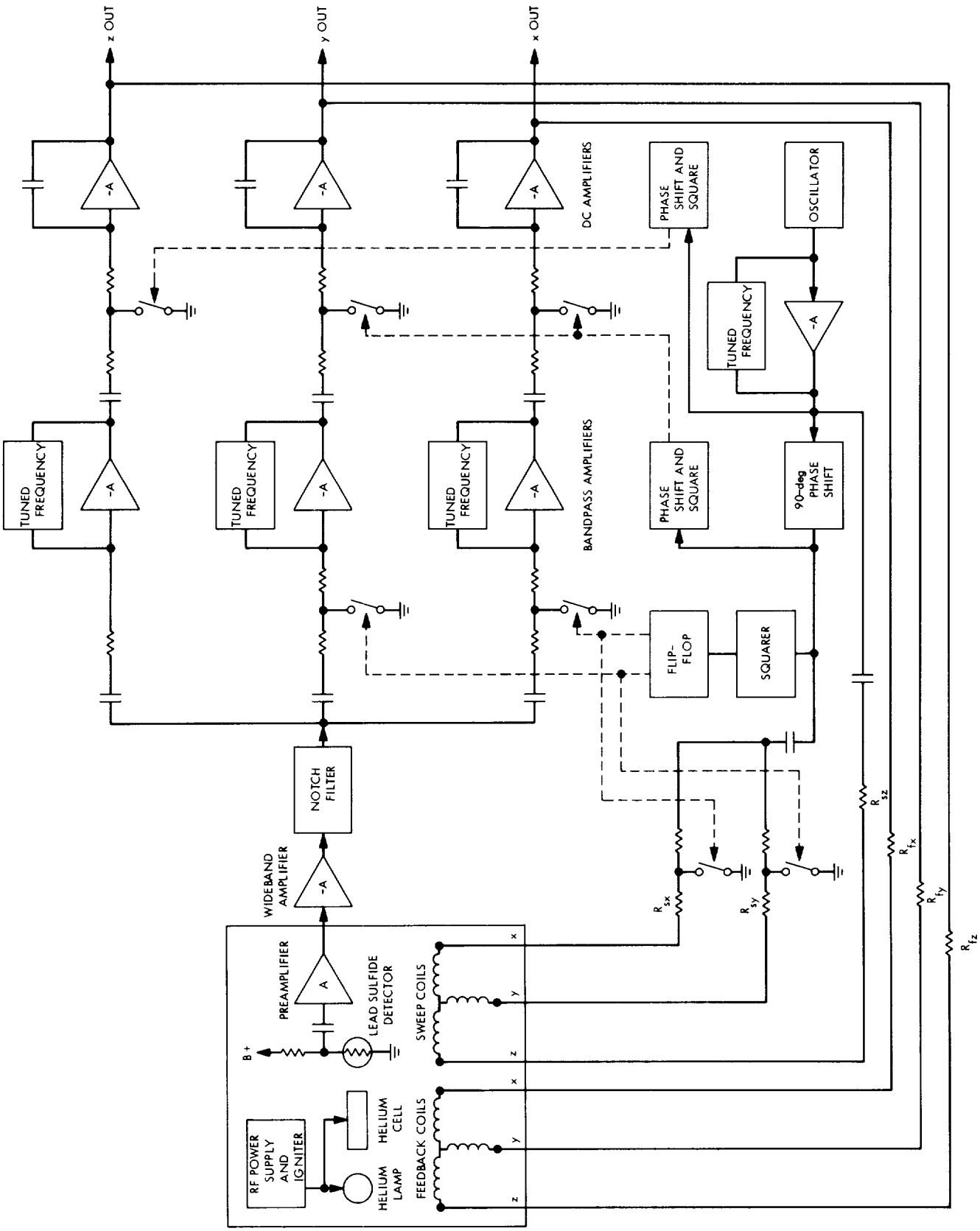


Fig. 1. Advanced helium vector magnetometer

two axes at one time. By commutating between xz and yz rotating vectors, a triaxial system is obtained. An oscillator generates a signal that is made a pure sinusoid by a bandpass amplifier. This drives the z-axis sweep coil with a sinusoidal current and provides the z-axis demodulator drive. The oscillator output is also phase-shifted 90 deg to provide the x and y sweep currents and demodulator drives, which are commutated between the x and y axes on alternate cycles. The signal from the sensor is then amplified and filtered to eliminate the second harmonic of the sweep vector. This signal is fed into the z-axis bandpass amplifier, commutated, and fed into the x- and y-axes bandpass amplifiers. The signal is then synchronously demodulated with respect to the corresponding component of the sweep vector to yield a dc signal that drives the dc amplifiers. Each dc output is then applied to the appropriate feedback resistor R_f which is in series with a sensor coil. The current through the respective x, y, and z coils produce a magnetic field to buck out the ambient field. Because the field of the coil is linear with applied current, the x-, y-, and z-axis output voltages are directly proportional to the associated fields, and are used as a measure of the magnitude of the ambient magnetic field.

Prior to this study, it was assumed that the helium vector magnetometer sensor would not perform in accordance with Eq. (3) for magnetic fields above approximately 500 γ. However, by increasing the magnitude of the sweep vector to correspond to the full-scale value of the feedback system, operations in accordance with Eq. (3) have been observed up to 100,000 γ (1 G), the limit imposed by present electronics. This is believed to be the first time that magnetic fields above 1000 γ have been measured vectorially using the optically pumped helium.

The magnitude of the sweep vector for the three axes is determined by the sensor coil constant, the output voltage of the sweep circuitry, and the sweep resistor R_s in series. The scale factor of the instrument is determined by the coil constant, the output voltage range, and the feedback resistor R_f. To maintain a constant-loop gain, the open- and closed-loop gains should vary directly. As indicated in Eq. (3), the sensor response to a given magnetic field decreases inversely with increasing sweep vector magnitude. Figure 2 shows the theoretical and actual sensor response to changes in H/H_s, given in terms of system ac gain.

Using the system shown in Fig. 3, and assuming infinite input and zero output amplifier impedances, the closed-

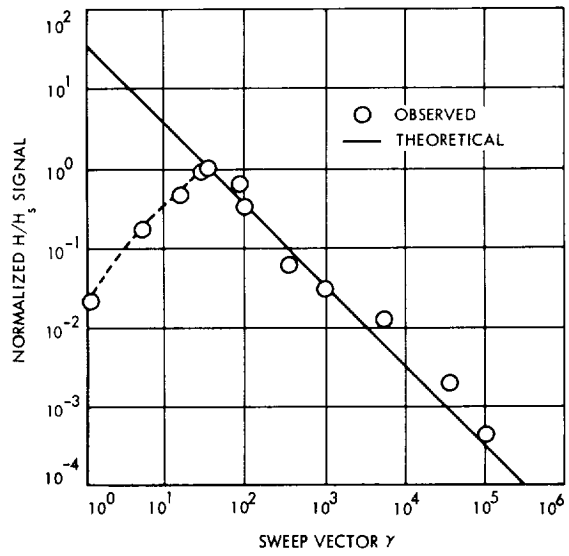


Fig. 2. H/H_s signal vs sweep vector

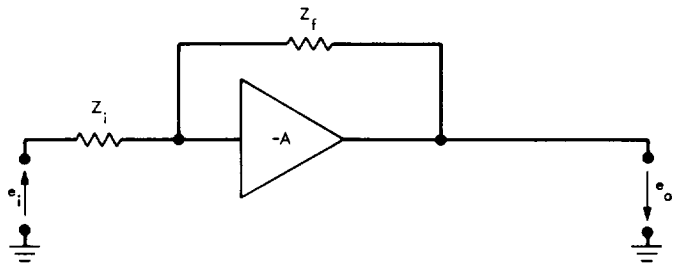


Fig. 3. Basic feedback configuration

loop gain is

$$\frac{e_o}{e_i} = -\left(\frac{Z_f}{Z_i}\right)\left[\frac{1}{1+\left(\frac{1}{A}\right)\left(1+\frac{Z_f}{Z_i}\right)}\right] \qquad (4)$$

where

e_i = input voltage

e_o = output voltage

Z_f = circuit feedback impedance

Z_i = circuit input impedance

A = amplifier gain

Letting $1/\beta = 1 + (Z_f/Z_i)$, where β is the feedback attenuation, the error term then becomes $1/[1 + (1/A\beta)]$ and the percentage error is a result of the product $A\beta$.

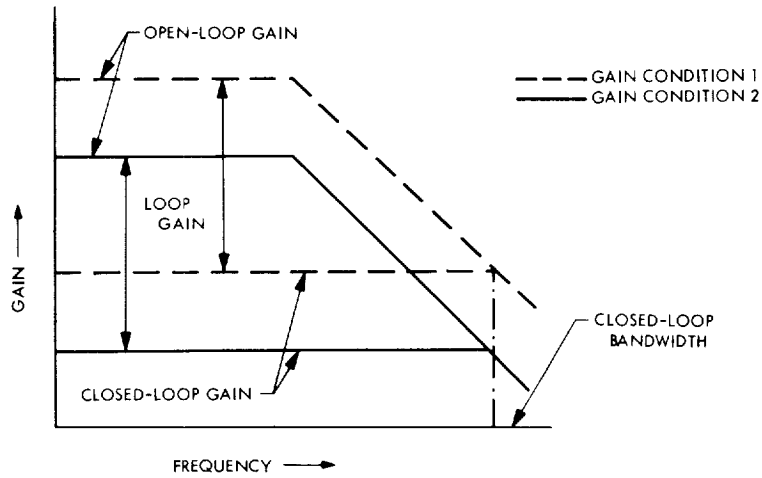


Fig. 4. Open- and closed-loop gain vs frequency for constant-loop gain

Figure 4 indicates how maintaining a constant-loop gain in this system, where the open-loop breakpoint is independent of gain (due to the output integrator), yields a fixed breakpoint in the closed-loop response. Therefore, the closed-loop system bandwidth remains constant as long as the loop gain remains constant.

In the helium magnetometer, the sweep vector is approximately equal to the full scale factor for optimum performance. This means that as H_s is increased, the loop gain is decreased directly. Therefore, as β goes up, A goes down, and vice versa, maintaining a constant $A\beta$ term. Some deviations from this are experienced at low fields, where the H/H_s relationship no longer holds true, but the portion of the dynamic range over which this occurs is small indeed, as seen in Fig. 2, and is not of important consequence over the dynamic range being considered.

A simple change of the feedback resistors R_f and the sweep field determining resistors R_s will then yield a magnetometer with a dynamic range of at least 10^6.

It has been demonstrated that the helium magnetometer performs well at both high and low fields. This discovery should greatly expand the areas of possible application of the helium vector magnetometer. A wide dynamic range instrument is possible using extremely simple range switching. An important use of the high-field operation for low-field magnetometry is the capability of isolating the instrument and magnetic field noises. This assurance is not present when the unit is operating as a low-field magnetometer, since shielding from the ambient field is incomplete.

Reference

1. Colegrave, F. D., and Franken, P. A., "Optical Pumping of Helium in the $^3S_1$ Metastable State," *Phys. Rev.*, Vol. 119, No. 2, pp. 680–690, July 15, 1960.

B. Ground Support System for X-Ray/ Gamma-Ray Prototype Flight Experiment,

L. L. Lewyn

A ground support system for a prototype flight pulse-height analyzer has been constructed and tested. The ground support system provides both command and readout capability for the flight instrument, as well as power and signal monitoring functions.

The flight data consists of pulse-height analysis and event scalar words that are transmitted in a serial bit stream with appropriate frame synchronization and parity checks. The pulse-height analysis words are decommutated and presented in parallel to a Nuclear Data ND 130A pulse-height analysis system for storage. The ND 130A is read visually on an x–y oscilloscope and data printout is provided by an IBM typewriter. Event scalar words appear on a Nixie visual readout and are printed out on a Hewlett-Packard HP 563A parallel printer. The printout also contains three digits of frame identification to provide time correlation. Parity and frame synchronization are indicated visually by Nixie readout. Pulse-height analysis words with parity errors are stored in half of the ND 130A memory, which is reserved for that purpose.

The flight system requires two voltage level commands and two pulse commands. The level commands control

the power and veto functions in the flight instrument and are actuated by switches on the ground support equipment (GSE) panel. The pulse commands step the instrument mode and the high-voltage power-supply programmer. The pulse commands are actuated by push-buttons on the GSE panel. All commands are buffered by integrated circuit logic.

Wideband current monitoring of the flight system is provided by oscilloscope inspection of the voltage differential across a 1-Ω series resistance. Flight unit operating time is monitored by a running time meter which is controlled by the power command switch. Monitoring is provided for four of the flight system signals that are routed to BNC connectors on the GSE front panel.

The GSE is constructed using discrete components and Amelco 300CJ-series high-noise immunity logic. Approximately 100 discrete components and 90 logic flatpacs are mounted on 9 general-purpose printed-circuit cards designed for this system by JPL.

The GSE will be used to support prototype flight system tests, provide a data acquisition facility for field evaluation tests of X-ray and gamma-ray detectors, and serve as a prototype for future GSE systems.

C. Parallel to Serial Converter for Punch Tape Perforation, *L. L. Lewyn*

A parallel to serial converter and punch tape drive has been completed and is in operation in the laboratory. The device is used for readout of X-ray and gamma-ray spectra. The use of the perforator eliminates the previous requirement for manually preparing punch cards from pulse-height analyzer printer output data. The punch tape output of the converter is fed directly into a PDP 4 which formats the data on magnetic tape for further reduction by an IBM 7094.

The parallel to serial converter accepts data from Victoreen ST 400 D and ST 800 DM pulse-height analyzers and presents the data one decade at a time to a Tally 1505 punch. Data readout begins with the most significant decade of pulse-height analyzer address zero. A 4.5-ms punch tape feed signal occurs first. The punch tape feed signal energizes the punch capstan solenoid. During the next 4.5-ms interval, the sprocket and data punch signals occur. The capstan is still rotating while the sprocket and data punch levers begin to move. However, by the time the punch levers strike the tape surface, the capstan rotation has stopped and the tape is no longer in motion. The levers punch the state of the most significant decade of pulse-height analyzer data into the tape in ASC II format. Odd lateral parity is also recorded. The feed and data punch signals are repeated for a period of 18 ms until all 6 decades of the first address are read out. During the seventh interval a space is punched. Readout then resumes with the most significant decade of the next address and continues through the remaining addresses.

The converter uses discrete components for pulse-height analyzer interface circuits and punch solenoid drive. Motorola MC 830P-series dual in-line integrated circuits are used to implement the logic functions. The device is constructed in a 19-in. card file for rack-mounting with the perforator (Fig. 5). The front panel controls allow leader and identification data to be manually punched into the tape.

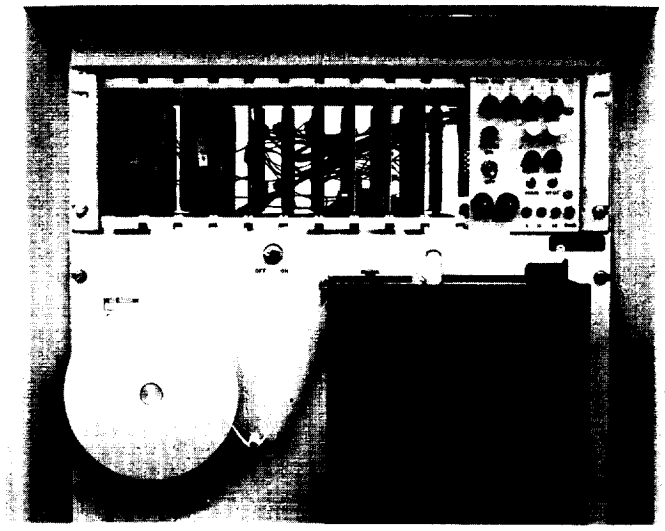


Fig. 5. Serial converter and punch tape perforator configuration

D. Electronics System for Measuring Induced Photon Spectra at the UCLRL Bevatron, *L. L. Lewyn*

An electronics system has been constructed for measuring induced photon spectra at the University of California's Lawrence Radiation Laboratory (UCLRL). The principal function of the system is the acquisition of 100 keV to 10 MeV gamma-ray spectra generated by a flux of high-energy protons on a thick target. The system also provides gating, timing, and monitoring functions.

The system is housed in an experiment shed located near the outer shield wall at the northwest section of the Bevatron. On the left side of the experiment shed, equipment made available by the Alvarez Physics Group of the Lawrence Radiation Laboratory is mounted on three racks. One rack contains the high-voltage supplies for the coincidence scintillator–photomultipliers, the second the high-speed counters, and the third the timing electronics. Located on the right side of the experiment shed are the high-speed oscilloscope in the magnetic shield (made available by the Alvarez Physics Group), the magnet power supply monitor and control system (provided by the Bevatron), the signal distribution panels, the delay line amplifiers, the gamma-ray detector high-voltage power supplies, the main resistance–capacitance amplifier, the monitor scope, the IBM output writer, and the resistance–capacitance–inductance (RCL) pulse-height analyzer (PHA).

The beam monitor and gamma-ray detector block diagram is shown in Fig. 6. The circulating beam is extracted by means of the north-outside-west target which is plunged into the beam for a preprogrammed interval that ranges from 400 to 800 ms. While the target is in the beam, the scattered protons are deflected by C magnet 564. The protons are then focused by quadripole magnet 562. After emerging from the Bevatron wall, the protons strike thin plastic timing scintillators S1 and S2, which are approximately 6 in. apart. The phototubes S2A and S2B collect the light from scintillator S2. The protons are deflected by Apollo magnet 567 and pass through closely spaced scintillators S3 and S4 before striking the target.

The 3 × 3-in. NaI (Tl) gamma-ray scintillator–photomultiplier is in close proximity to the target. A charge-sensitive preamplifier located a few feet from this detector performs a short time constant charge-to-voltage conversion on the photomultiplier signal and drives the signal over approximately 70 ft of cable from the target area to the experiment shed. Three layers of magnetic shielding were required around this detector to prevent photopeak shifts during the high Bevatron magnetic fields at beam spill time. The monitor detector is located approximately 10 ft from the beam line and consists of a 3×3-in. NaI (Tl) scintillator–photomultiplier ray detector. The monitor detector signals are used to obtain gamma-ray background levels.

The timing electronics block diagram is shown in Fig. 7. Photomultiplier signals S1A through S4B are passed through variable electronic delays. The delays are

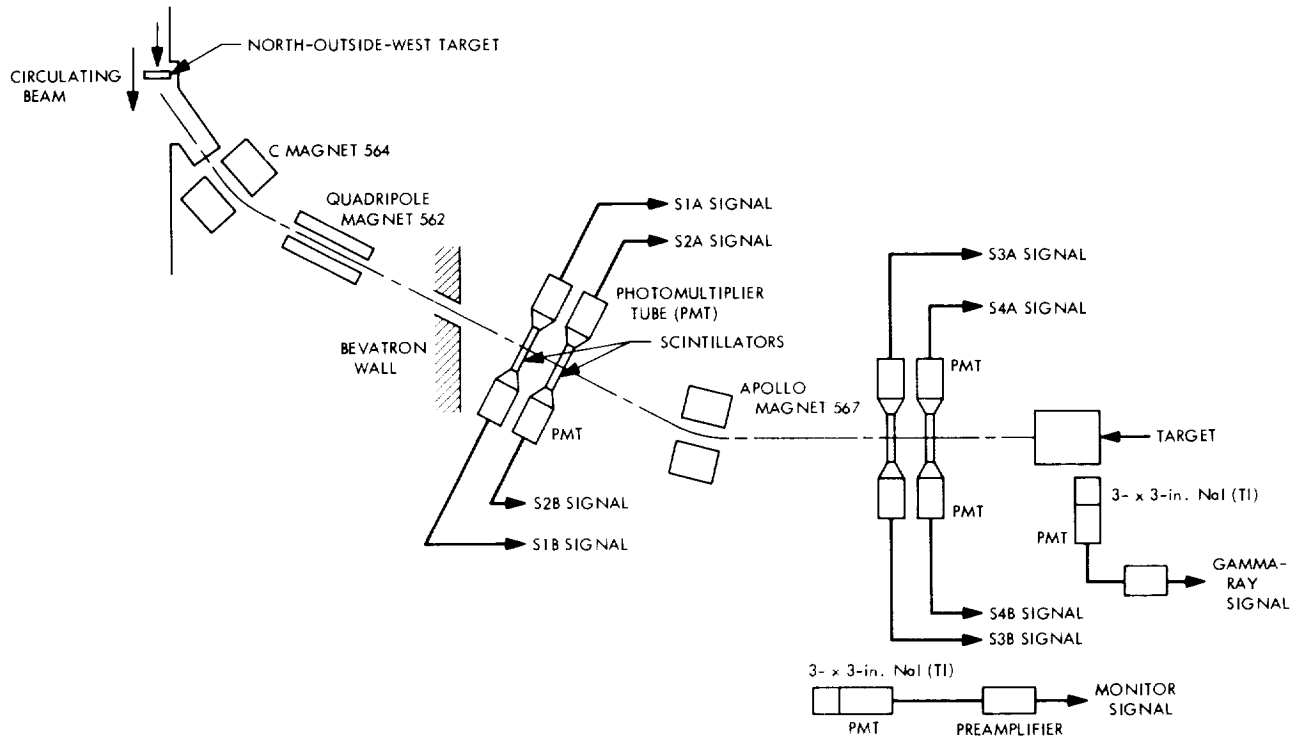


Fig. 6. Beam monitor and gamma-ray detector block diagram

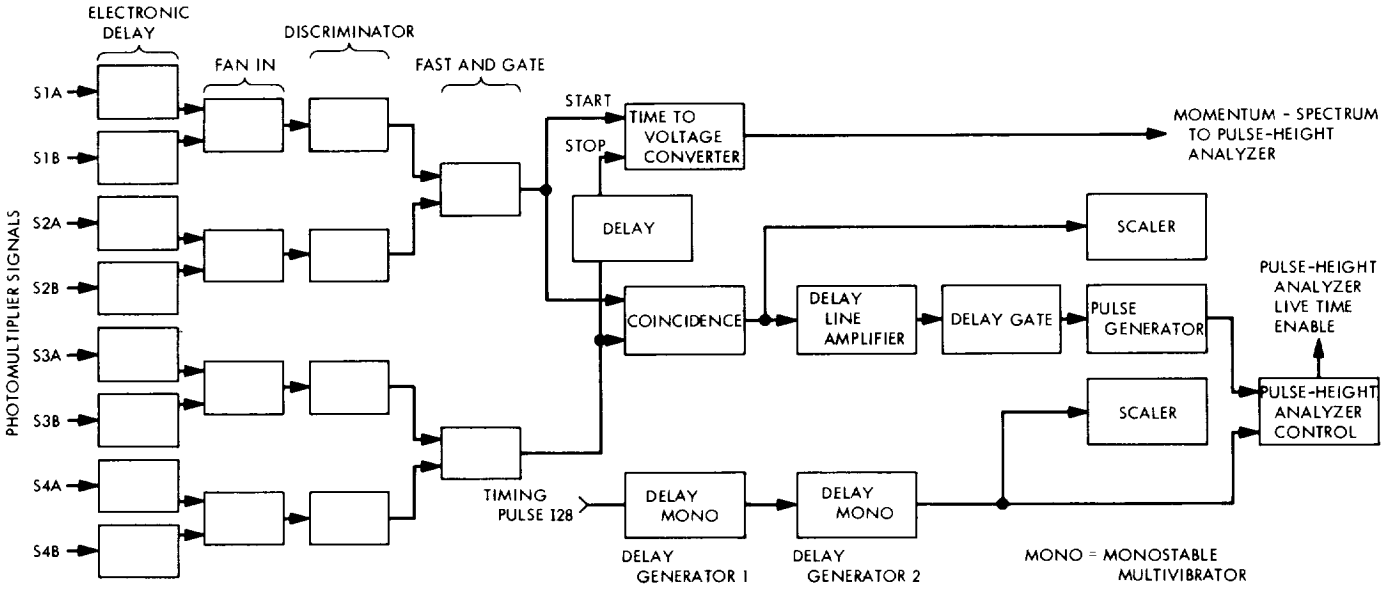


Fig. 7. Timing electronics block diagram

adjusted so that all signals resulting from the passage of a proton in the energy range of interest arrive simultaneously at the fan-in networks. Proper output signals from the fast *and* gates should then arrive simultaneously at the input of the coincidence network. The coincidence network has a resolving time of 7 ns. The coincidence network output is monitored by a high-speed scaler, which records the flight of each proton within the energy range of interest. The low-level output of the coincidence network is amplified and delayed. The delayed output triggers a pulse generator, which provides an acceptance gate input level to the PHA control for the duration of the pulse. The presence of another signal is also required for the PHA control to allow the analyzer to store. This signal is present for the duration of the spill and is derived from the Bevatron I28 timing pulse. The I28 pulse triggers two delay generators in succession. The first generator output rises at I28 and lasts until the start of the spill. The trailing edge triggers the second generator. The output of the second generator is present during the spill. For many of the runs the output of the delay generator was used to turn on the PHA control without requiring the presence of the coincidence pulses.

Beam momentum analysis by the time-of-flight method was obtained by using the coincidence network inputs to drive a time-to-voltage converter. The coincidence network input from the first scintillator pair (S1–S2) is connected to the *start* input of the time-to-voltage converter. The coincidence network input from the second scintillator pair (S3–S4) is connected to the *stop* input. Electronic delays have been added to the S1–S2 signal chain so that signals from a proton in the middle of the energy range of interest will not arrive at the coincidence

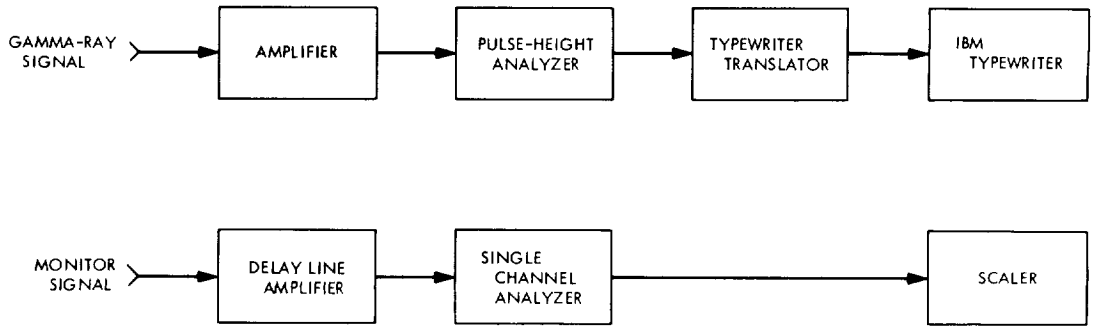


Fig. 8. Gamma-ray signal analysis block diagram

network simultaneously and have an apparent zero time-of-flight. Therefore, a delay comparable to the S1–S2 delay is added to the *stop* input of the time-to-voltage converter to center the time-of-flight spectrum peak above a zero value.

The gamma-ray signal analysis block diagram is shown in Fig. 8. The gamma-ray signal from the preamplifier in the target area is amplified by a TC 200 amplifier and routed to the low-voltage input of an RCL pulse-height analyzer. The gamma-ray spectra accumulated in the PHA are printed out on an IBM typewriter. The monitor input is amplified by a delay line amplifier and routed to a single-channel analyzer with window levels set to 300 keV and 3.0 MeV. The single-channel analyzer output is monitored by a scaler to provide a measure of gamma-ray background levels.

Experiments at the northwest section of the Bevatron have been completed and the electronics system is in storage. In late 1968, experiments are expected to be resumed in the external beam area.

E. Developmental Flight-Model Infrared Interferometer, *R. A. Schindler*

1. Introduction

The objective of the flight-model infrared interferometer program is to develop a flightworthy, high-resolution, infrared interference spectrometer for analysis of the earth's atmosphere and for surface and atmospheric analysis on advanced planetary missions. The instrument is to be used as a high-resolution infrared spectrometer for measuring the absolute intensity of spectral emission or absorption lines in the frequency interval 2000–8000 cm^{-1}. The design resolution is 0.5 cm^{-1}.

The operation, functional components, and design considerations of this instrument were presented in SPS 37-43, Vol. IV, pp. 253–257; the basic operating principles, as well as some results of Fourier spectroscopy, were described in SPS 37-30, Vol. IV, pp. 164–171. Progress in the design, development, and construction of this instrument is discussed in the following functional areas:

(1) The optical portion consisting of: (a) the basic interferometer, including the cat's-eye retroreflector and beam-splitter cube; and (b) the fore-optics, including the energy collection mirrors and optical chopper.

(2) The cat's-eye position stabilizing and stepping system (servo-drive).

(3) The data acquisition system consisting of the infrared detection system and the telemetry system.

The experiments to be conducted with this instrument and the instrument requirements imposed by these experiments will be reported in a separate article.

2. Basic Optical Portion of the Interferometer

Since the last report, two cat's-eye retroreflectors (Fig. 9) have been successfully assembled and focused. Tests have shown their performance to be satisfactory for the successful operation of this instrument.

No real problems were encountered after assembly except for a quartz secondary mirror support, which cracked after assembly. The problem was solved by making these supports thicker. It is difficult to fasten these supports to the quartz tube, which positions the secondary mirror at the focal point of the primary mirror, because the optical alignment of the secondary mirror-driving transducer-support assembly (cemented together beforehand) must be accurate before the support is cemented to the quartz tube. This alignment is aided by a small hole in the center of the primary mirror that enables the secondary mirror to be viewed with a telescope.

Rough adjustment of the primary-to-secondary distance was initially done with the vibrating Ronchi ruling method (SPS 37-43, Vol. IV). Fine adjustment was done with the single cat's-eye interferometer arrangement shown in Fig. 10, which uses a photomultiplier to indicate the amplitude of the interference fringes as the differential thread adjustment on the cat's eye is moved. The cat's eye is in correct adjustment when the fringe amplitude is a maximum. It was found that the Ronchi method was not sensitive enough for the final adjustment. In fact, the rough adjustment can be done just as well with the interferometric method by visual observation of the fringes.

For the future, it is intended to design an all-quartz cat's eye. There will be no mechanical adjustment and the unit will be fabricated to the correct optical tolerance by trial-and-error methods. This cat's eye will be more stable and the weight will be reduced by a factor of about 3 from the 2 lb of the current units. The reduction in mass is not particularly important in itself. However, the reduction in inertia will allow the size of the magnet and drive-coil to be reduced and, consequently, the weight of the entire structure will be reduced.

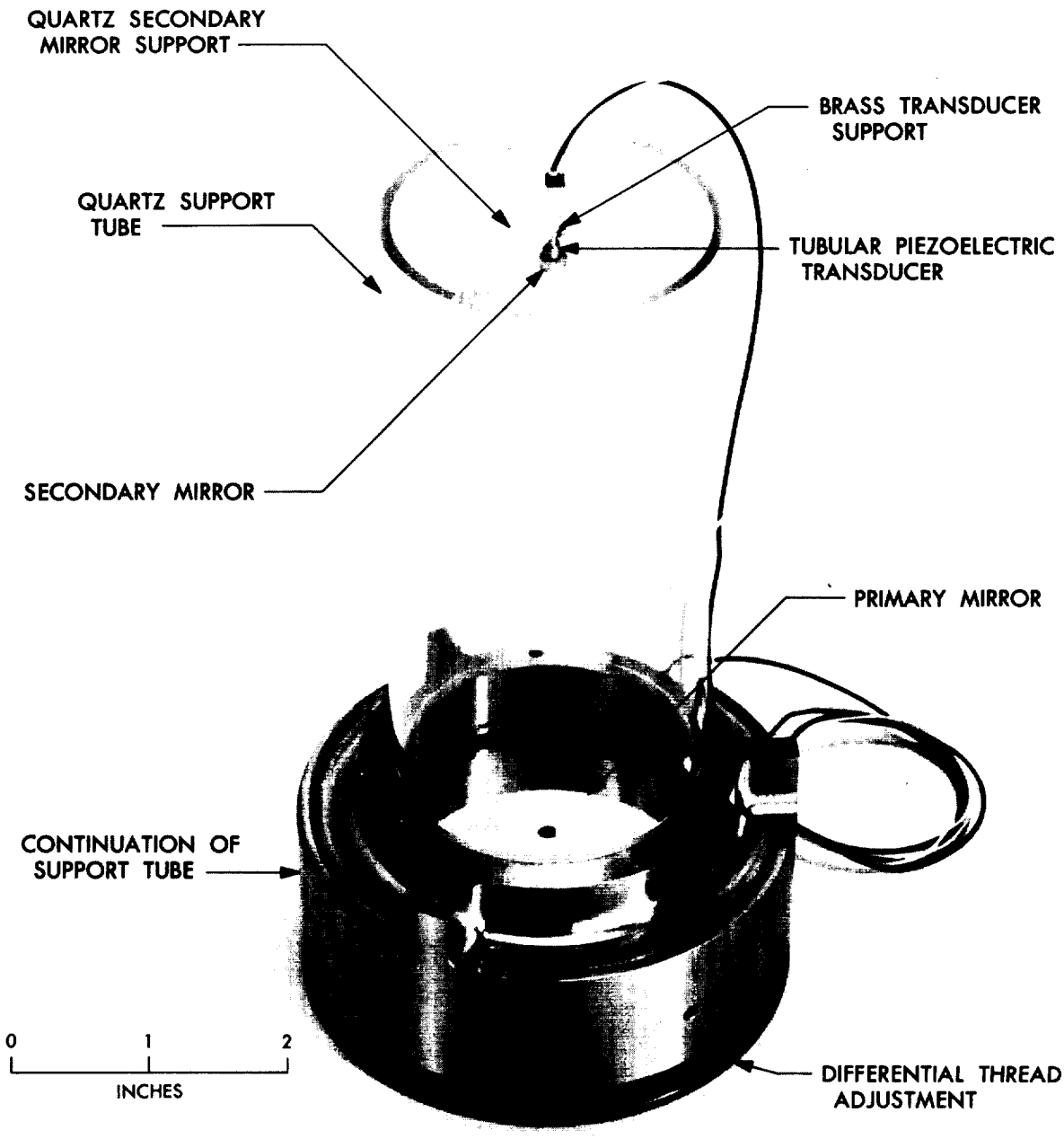


Fig. 9. Cat's-eye retroreflector

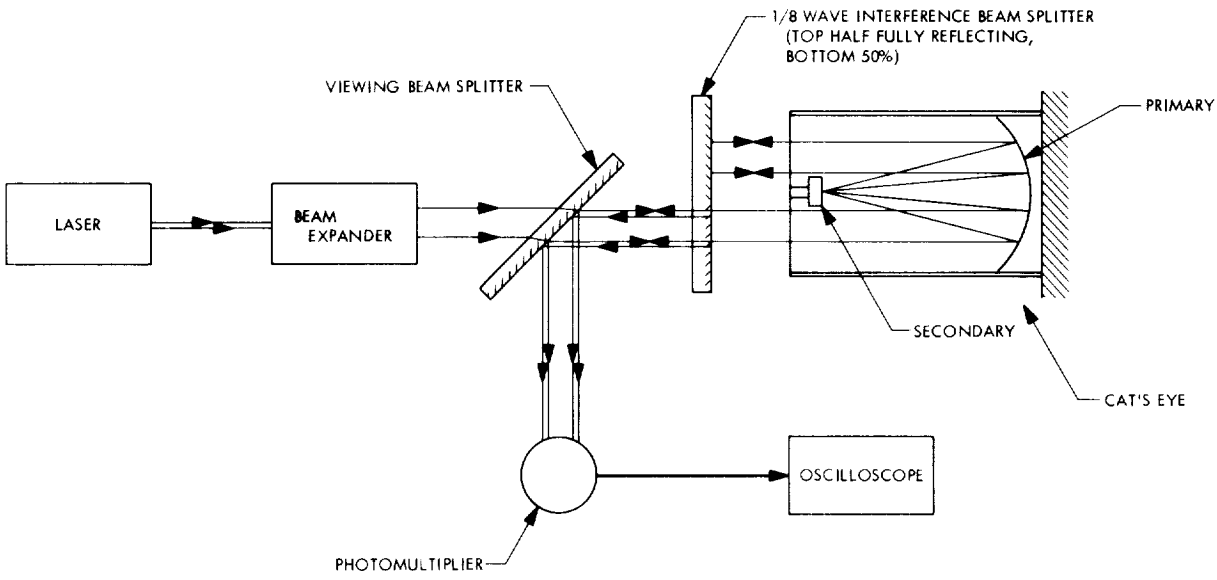


Fig. 10. Interference focusing arrangement

Although a quartz beam-splitter cube is available for the interferometer, which will suffice for testing the servo-drive system, an infrared transmitting beam splitter is required for the final system. Such a beam splitter (Fig. 11), capable of transmitting radiation to 5 μm, has been designed and is being manufactured[1] out of single-crystal calcium fluoride. Unnecessary portions of the cube have been removed, which should increase the infrared transmission and may also reduce distortion effects caused by variations in the index of refraction that are greater with calcium fluoride than with glass. The angles A and B (Fig. 11) may not differ by more than 0.2 arc sec if the system is to operate as an axially symmetric interferometer. The beam-splitting surface for the infrared portion will be a ¼-wave (at 2 μm) coating of ultrapure silicon while the reference channel will be partially silvered.

[1]F. J. Cooke, Inc., North Brookfield, Mass.

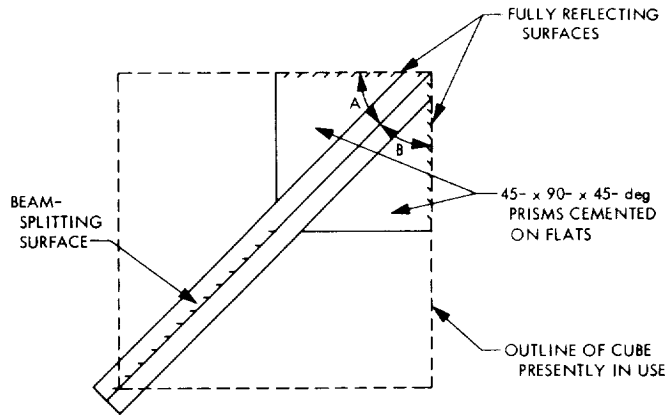


Fig. 11. Beam splitter

3. Fore-optics

The fore-optics (Fig. 12) is that part of the optical train which accepts the incoming radiation, impresses a modulation on it, and introduces a comparison source. This combined radiation is collimated and fed to both the interferometer and a reference, total-power detector.

Difficulty in procuring an adequate optical tuning-fork chopper was anticipated because of the requirement that the reflected image from the blade (calcium fluoride) have an angular stability of ± 1 arc min for the required motion amplitude of 1.2 mm. However, a chopper meeting these requirements has been produced.[2]

The 90-deg off-axis gold-coated paraboloid mirrors (focal length 50 mm) have been received and appear to be satisfactory. These mirrors are used to condense the radiation so it can be chopped, recollimated for the interferometer, and finally condensed on the infrared detectors.

The optical assembly for the servo-drive reference source, a mercury lamp, has been completed and was

[2]Manufactured by American Time Products Div., Bulova Watch Co., Inc., Woodside, N.Y.

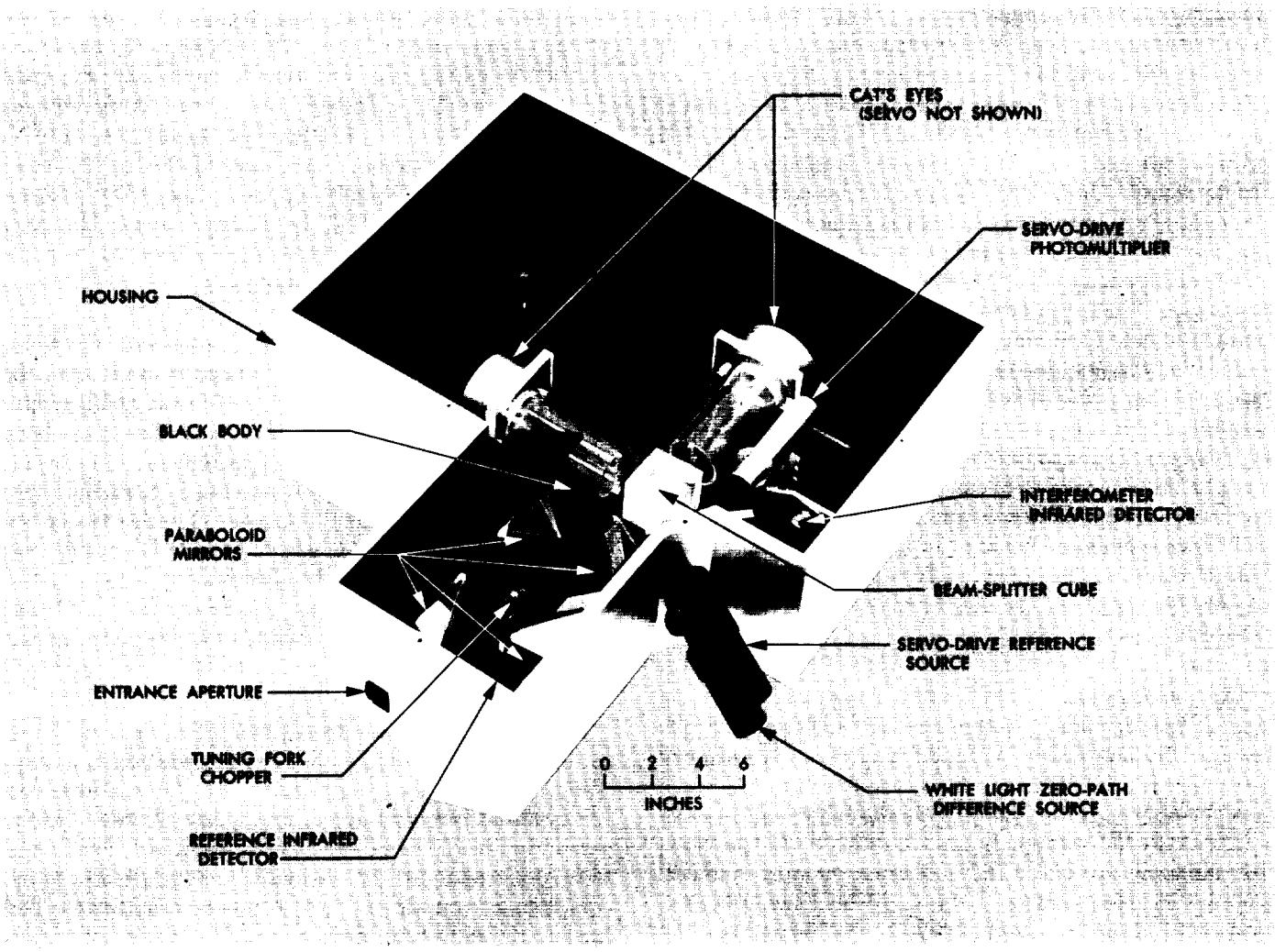


Fig. 12. Interferometer model.

utilized in the servo-drive test. A laser reference source, using a 9-in.-long helium–neon (6328 Å) laser tube, is now being designed because a laser reference will give better servo performance due to its much higher intensity. This laser will be enclosed in a hermetically sealed, pressurized box, along with its power supply, to eliminate high-voltage corona problems inherent at balloon flight altitudes. This packaging approach is also being used with the servo-drive photomultiplier tube.

Design of the fore-optics housing and mirror mounts has been completed. The fore-optics housing was fabricated as a single unit using electron-beam welding to prevent stresses and consequent warpage during machining operations. This housing contains the cat's eyes, beam splitter, and servo-drive assembly, as well as the fore-optics and detectors. It was built out of ⅝-in.-thick aluminum plate so that it would be sufficiently rigid to eliminate resonances which could be detrimental to the operation of the interferometer. The weight is presently about 81 lb, but, after successful operation of the interferometer system has been demonstrated, the weight can be reduced by a factor of 3 by milling away unnecessary material.

The mirror mounts were fabricated by electrolytic-discharge machining because conventional machining methods could not have economically produced the cavities capable of firmly holding these odd-shaped mirrors.

4. Cat's-Eye Position Stabilizing and Stepping System (Servo-Drive)

The function of the servo-drive is to move the cat's eye in equal steps, whose size is one wavelength of the

monochromatic reference source. Some changes have been made to the basic operation of the system. One of these is the addition of a reversible fringe counter (Fig. 13). Two signals in phase quadrature are required to operate a reversible counter: one is the demodulated error signal and the other is the dc component of the fringe signal. These two signals can be seen to be in phase quadrature by considering the demodulated signal, obtained by vibrating the path difference a small fraction of one fringe to be the first derivative of the dc level of the fringe. Since the fringe amplitude from the monochromatic reference source vary sinusoidally with path difference, the first derivative also varies sinusoidally but has a 90-deg phase difference. The function of this reversible counter is to keep track of the actual number of fringes traversed, regardless of the direction of motion. Thus, if the servo were to temporarily fail for some reason (such as too high an acceleration), the exact optical path difference would still be known. Knowledge of the exact path difference is of utmost importance because an error of only one step position can invalidate the entire interferogram.

The command to step to the next fringe is given to the system incrementing another counter (reference fringe counter) by one count. A digital magnitude comparator compares the numbers on the reversible and reference fringe counters. If the numbers are the same, the cat's eye will stay in the present fringe position. If the numbers are different (as when the reference counter is incremented by one count or if the servo becomes unlocked and is in error by many fringes), the servo stepping control is commanded to restore the digital balance. The number on the reversible counter is telemetered. In addition, a flag to indicate that the numbers are not the same is sent. Originally, it was proposed to drive the secondary mirror vibrator, which modulates the path difference, passively. Instead, the mirror, along with its piezoelectric driver, now forms the resonant element in a crystal oscillator circuit. Tests showed this to be necessary to obtain a sufficiently high amplitude (20 Å peak-to-peak) without overheating the piezo element. The resonant frequency is 480 kHz, which is high enough to allow for a 20-kHz servo-system bandwidth.

A laboratory test was made to demonstrate the feasibility of the basic servo-control system. In particular, the use of the secondary mirror to provide the small amplitude, high-frequency error corrections had not been demonstrated before. The test arrangement utilized portions of an interferometer used previously to demonstrate the feasibility of the "double-passed" interferometer approach. A simple actuator was built and laboratory test equipment was used for the electronic portion. In this test, the path-length vibrator secondary was driven at 150 kHz; the system bandwidth was limited by the phase-demodulator instrument to 1 kHz.

Due to unsatisfactory progress in the design and fabrication of the servo-drive, the vendor[3] was given a stop-work order in September 1967 and the contract was renegotiated. Work resumed on January 12, 1968 and the assembly was delivered on June 21, 1968. JPL acceptance tests performed at the vendor's facility demonstrated compliance with all but two of the specifications called forth in the statement of work, but these were not serious

[3]Aeroflex Laboratories, Plainview, N.Y.

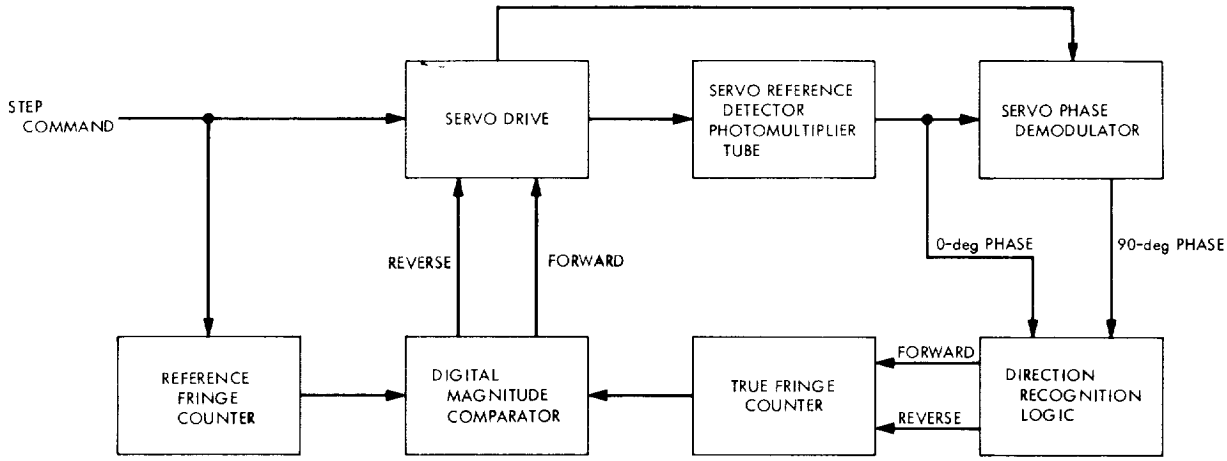


Fig. 13. Fringe counters and direction command logic

enough to affect operation of the instrument. The length of travel of the moving cat's eye was slightly less than the 1 cm called for but an adequate tolerance had been allowed. Coil resistance and number of turns was slightly out of specification but will have no effect on operation.

Completion of the electronic design and construction is now an in-house effort. To date, all circuit boards have been fabricated and tested but systems testing awaits fabrication of card files, cables, etc.

5. Data Acquisition System

a. Infrared detection system. The infrared detection system (Fig. 14) includes the infrared detectors, ac amplifiers for amplifying the chopped signal, demodulator, integrator, and analog-to-digital converter (ADC). Considerable changes have been made since the last system report.

Signal amplification is done in two steps. Adjacent to the infrared detector is a high input–impedance preamplifier utilizing a field-effect-transistor input operational amplifier. The gain can be varied in steps, upon command, to a set of reed relays which decrease the amount of feedback. During a flight, the gain changes will be preprogrammed according to the number of reference fringes from the zero-path-difference position because the amplitude is large only within a few tens of fringes around the zero-path difference. Varying the feedback for the gain changes, instead of using an input attenuator, provides maximum system linearity as well as a better noise figure.

The preamplifier is followed by another ac amplifier which also uses an operational amplifier and feedback controlled-gain changes. The gain of this amplifier varies inversely as the signal integration period, allowing for signal-level equalization at different stepping rates and integration times.

The demodulator must be capable of asynchronous operation because of the start–stop mode of operation of the servo system. In the fastest mode of operation (500 fringe steps/s), the detector will see the radiation source and reference blackbody only once each step (i.e., for a total time of 1 ms). The radiation from the interferometer

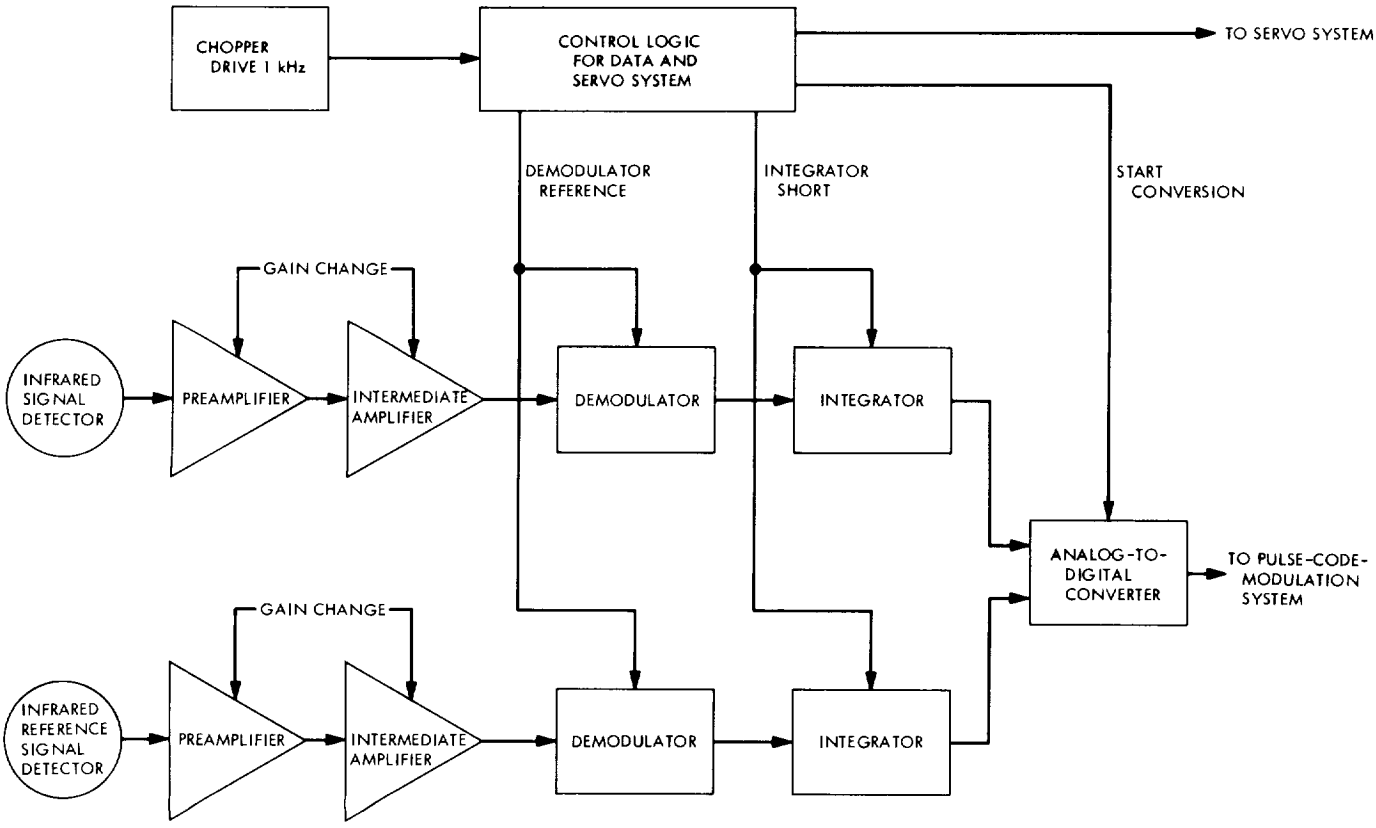


Fig. 14. Data-system block diagram

can be utilized only during 1 of the 2 ms available because the servo is stepping during the other. When the interferometer is being stepped at slower rates, however, more time is available for signal integration (e.g., at 250 fringe steps/s, 1 ms is still lost during stepping but now 3 ms are available for integration). The demodulator must not have a "memory" of the previous data point because the signal can vary greatly from one step to the next.

The signal integrator, which follows the demodulator, is also designed for asynchronous operation. First, it integrates the demodulated signal for an integral number of chopper cycles (e.g., 3 in the 250-step/s case given above), then it holds the integrated signal for the first half of the 1-ms stepping period to allow the ADC to digitize the signal. This method eliminates the necessity for a separate sample-and-hold circuit. During the second half of this 1 ms, the integrating capacitor is shorted by a transistor switch to set the integrator back to zero.

The ADC is of the successive approximation type wherein reference voltages, decreasing in steps of two, are compared to the signal and are left on or turned off depending upon whether they are less or greater than the signal. No other type of converter can digitize rapidly enough for the required resolution. The circuits have been built on standard module cards. The in-house-designed circuits are built on blank cards of the same type as the ready-made logic cards. Such construction methods are suitable for balloon flights, for which the engineering model has been designed.

b. Telemetry system. Design of the balloon flight telemetry system is now underway. It is presently planned to build the data-encoder portion of the pulse-code-modulation (PCM) telemetry system in-house because of the unique synchronization problems with the optical chopper and stepping rate of the interferometer. As presently planned, there will be around 50 bits/word/step, which includes the infrared data, the fringe number, a flag if the data point is no good, and the detection system gain settings. Every sixteenth data word will give the reference detector intensity instead of the fringe intensity; the interferometer will not step at this time so that no fringes will be missed. The interferometer will also not step during the frame synchronization time for the same reason. The total bit rate at 500 steps/s will be about 21,500 bits/s and will be proportionately slower at slower step rates. The same PCM format will be adhered to regardless of the step rate, simplifying the ground data-handling problem.

The logic circuitry for the PCM data encoder will, as in the other areas, be built out of standard modular plug-in cards using diode-transistor logic. The ground data-handling system, on the other hand, will be assembled out of commercially built equipment. It is intended to use a small computer to decommutate the PCM data and format it to a standard 7- or 9-track IBM format magnetic tape so that the interferograms may be processed by an IBM 7094 or similar computer, using the same computer program presently in use for reducing planetary interferometer data.

An FM–FM telemetry system is planned for the balloon flights to simplify the transmission of engineering data and also to allow other experiments to operate simultaneously and independently of this experiment. The PCM data will be transmitted by a very wide-band ($\pm$66 kHz) subcarrier, which will be possible with S-band (2.2–2.3 kHz) telemetry. (By the time of the planned balloon flight, S-band telemetry will be mandatory.) FM–FM systems do not have as good a signal-to-noise ratio as simple FM systems, but this should not be a problem on a balloon flight because the signal level remains high as long as there is a line-of-sight to the balloon.

c. Laboratory data acquisition system. Because it will take considerable time to make a full PCM data-handling system operational, even without telemetry, an incremental 7-track IBM format stepping recorder will be used in the interim to obtain interferograms. The incremental recorder will record fringe number and data. When the recorder is ready for another data word, it will signal the servo-drive to step to the next reference fringe and the data will be integrated and digitized, and the next data word will be recorded. This maximum stepping rate possible with this scheme will be about 10 fringes/s. However, it should be possible to simulate the effect of high-speed operation (as far as the data system is concerned) because the integration time can be as small as necessary. Considerable preliminary testing will be possible with this recorder before the PCM system is completed.

XIV. Science Data Systems

SPACE SCIENCES DIVISION

A. Decomposition of the States of a Linear Feedback Shift Register Into Cycles of Equal Length, *M. Perlman*

1. Introduction

The linear logic feedback shift register (FSR) in Fig. 1 has been investigated in considerable detail in Ref. 1. The state of the ith two-state memory element at clock pulse interval (CPI) k is denoted as a_{k-i}. The behavior of the FSR can be described by the linear recurrence relationship

$$a_k = e + \sum_{i=1}^{r} c_i\, a_{k-i} \tag{1}$$

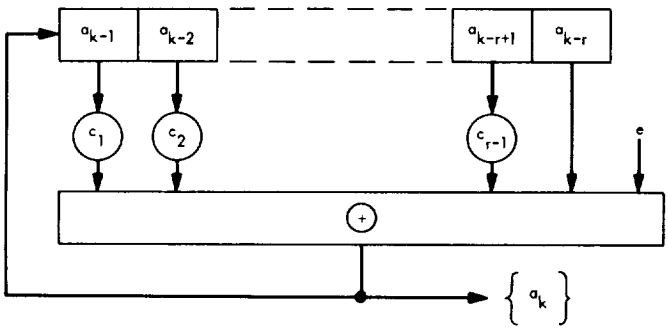


Fig. 1. Linear logic feedback shift register

The bit fed back at CPI k is denoted as a_k. The ith stage contributes to the feedback when the Boolean multiplier c_i is at state value 1. The summations are taken modulo 2 and e, a Boolean constant, is 0 for mod 2 summing (*exclusive-or*) or 1 for the complement of mod 2 summing (*not exclusive-or*). The feedback function in Eq. (1) is of the form to guarantee branchless cycles where distinct states have distinct successors.

The cycle length or periodicity of $\{a_k\}$ can be determined from its generating function (Ref. 1)

$$G_e(x) = \sum_{k=0}^{\infty} a_k\, x^k$$

$$= \sum_{k=0}^{\infty} \left[e + \sum_{i=1}^{r} c_i\, a_{k-i} \right] x^k$$

For $e = 0$

$$G_0(x) = \frac{\sum_{i=1}^{r} c_i x^i \left(a_{-i}x^{-i} + a_{-i+1}x^{-i+1} + \cdots + a_{-1}x^{-1}\right)}{1 + \sum_{i=1}^{r} c_i x^i}$$

$$= \frac{g_0(x)}{f_0(x)} \tag{2}$$

where a_{-i} is the initial state of the ith stage.

For $e = 1$

$$G_1(x) = \sum_{k=0}^{\infty}\left[1 + \sum_{i=1}^{r} c_i a_{k-i}\right]x^k$$

$$= \frac{1}{1+x} + \sum_{k=0}^{\infty}\sum_{i=1}^{r} c_i a_{k-i}x^k$$

$$= \frac{1}{x+1} + \sum_{i=1}^{r} c_i x^i \sum_{k=0}^{\infty} a_{k-i}x^{k-i}$$

$$= \frac{1}{x+1} + \sum_{i=1}^{r} c_i x^i \left(a_{-i}x^i + a_{-i+1}x^{i+1}\right.$$

$$\left. + \cdots a_{-1}x^{-1} + G_1(x)\right)$$

$$= \frac{\dfrac{1}{x+1} + g_0(x)}{f_0(x)} = \frac{1 + (x+1)\,g_0(x)}{(x+1)\,f_0(x)} = \frac{g_1(x)}{f_1(x)} \tag{3}$$

The characteristic polynomial $f_e(x)$ is a function of the feedback connections only. The length of the longest FSR cycle(s) is the smallest value of s for which $f_e(x)$ divides $x^s + 1$. The polynomial $g_e(x)$ is a function of the initial state of the register as well as the feedback connections. The degree of $g_e(x)$ is always less than that of $f_e(x)$. Whenever an initial state $a_{-1}\, a_{-2} \cdots a_{-r}$ yields a $g_e(x)$ that has a common factor with $f_e(x)$, it will lie on a cycle whose length divides s, the length of the longest cycle(s). Initial states which result in a $g_e(x)$ that is relatively prime to $f_e(x)$ will lie on a cycle whose length is equal to s. Thus all cycle lengths of a given FSR divide s.

2. FSR Configurations Which Yield Cycles of Equal Length

a. Characteristic polynomial. An r-stage linear FSR will decompose its 2^r states into cycles of equal length if it has the following characteristic polynomial:

$$f_1(x) = (x+1)\,(x+1)^r \tag{4}$$

The factor $x + 1$ in Eq. (4) is a result of complementary mod 2 summing. The actual feedback connections are determined from $f_0(x)$, which is equal to $(x+1)^r$.

First, the length of the longest possible cycle will be determined. Then it will be shown that no initial state can be selected which results in a $g_1(x)$ that has a common factor with $f_1(x)$.

b. Periodicity of $f_1(x)$. The length of the longest cycle is $s = 2^i$, where these inequalities are satisfied:

$$2^{i-1} < r + 1 \leq 2^i \tag{5}$$

Note that

$$(x+1)^2 = x^2 + 1$$

and by induction

$$(x+1)^{2^i} = x^{2^i} + 1$$

thus

$$f_1(x) = (x+1)^{r+1} \mid (x+1)^{2^i} = x^{2^i} + 1$$

The $00 \cdots 0$ initial state lies in the longest cycle, since this reduces $G_1(x)$ to $1/f_1(x)$.

Example 1. Given

$$f_1(x) = (x+1)\,(x+1)^6$$

$$= (x+1)\,(x^6 + x^4 + x^2 + 1)$$

where

$$a_k = 1 + a_{k-2} + a_{k-4} + a_{k-6}$$

$$r + 1 = 7$$

and

$$s = 2^3 = 8$$

The successive states for an initial state of 000000 is given in Table 1.

c. Structure and properties of $f_1(x)$. To determine the feedback connections, the $(x+1)^r$ factor of Eq. (4) must be expanded. Representing $x + 1$ as 11 and multiplying 11 by 11 gives 101 (using mod 2 arithmetic) or $x^2 + 1$.

Table 1. Successive states for initial state of 000000

k	a_{k-1}	a_{k-2}	a_{k-3}	a_{k-4}	a_{k-5}	a_{k-6}	a_k
0	0	0	0	0	0	0	1
1	1	0	0	0	0	0	1
2	1	1	0	0	0	0	0
3	0	1	1	0	0	0	0
4	0	0	1	1	0	0	0
5	0	0	0	1	1	0	0
6	0	0	0	0	1	1	0
7	0	0	0	0	0	1	0

This result is again multiplied by 11 resulting in 1111 or $x^3 + x^2 + x + 1$, the expansion of $(x + 1)^3$. Table 2 was generated in this manner. Expansion of $(x + 1)^r$ is tabulated for values of r from 1 through 16.

The expansion of $(x + 1)^r$ results in an even number of terms for all r. All the terms except the constant term 1 represent a feedback connection. Hence, the number of feedback connections is odd for all r. The number of feedback connections T may be expressed as follows:

$$T = 2^w - 1 \tag{6}$$

where w is the weight (i.e., the number of 1's) of the binary representation of r. The number T is a Mersenne number (Ref. 2).

Example 2. Given $r = 13$ or 1101_2 and $w = 3$,

$$(x + 1)^{13} = (x + 1)^8 (x + 1)^4 (x + 1)$$
$$= (x^8 + 1)(x^4 + 1)(x + 1)$$

Each term in the expansion has x^8 or 1 and x^4 or 1 and x or 1 as a factor. The expansion of $(x + 1)^{13}$ results in a total of 2^w or 8 terms and represents 7 feedback connections. Since the factors of $f_1(x)$ are self-reciprocal, $f_1(x)$ is a self-reciprocal polynomial. Thus each sequence $\{a_k\}$ associated with $f_1(x)$ in Eq. (4) and its reverse are identical. The self-reciprocity of $(x + 1)^r$ is apparent in the symmetry of the feedback connections. That is, c_i and c_{r-i} are equal for $1 \leq i < r$.

The recurrence relationship in Eq. (1) may be expressed as

$$a_k = h(a_{k-1}, a_{k-2}, \cdots, a_{k-r+1}) + a_{k-r}$$

where h is the *exclusive-or* or *not exclusive-or* Boolean functions of $r - 1$ or fewer variables. The recurrence relationships associated with $f_1(x)$ in Eq. (4) satisfy the following equation:

$$h(a_{k-1}, a_{k-2}, \cdots, a_{k-r+1}) = h(a'_{k-1}, a'_{k-2}, \cdots, a'_{k-r+1})$$

where

$$a'_{k-i} = 1 + a_{k-i}$$

Table 2. Feedback connections from the expansion of $(x + 1)^r$ for $r = 1, 2, \cdots, 16$

r in Binary	x^0	x^1	x^2	x^3	x^4	x^5	x^6	x^7	x^8	x^9	x^{10}	x^{11}	x^{12}	x^{13}	x^{14}	x^{15}	x^{16}	T
		c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	
1	1	1																1
10	1	0	1															1
11	1	1	1	1														3
100	1	0	0	0	1													1
101	1	1	0	0	1	1												3
110	1	0	1	0	1	0	1											3
111	1	1	1	1	1	1	1	1										7
1000	1	0	0	0	0	0	0	0	1									1
1001	1	1	0	0	0	0	0	0	1	1								3
1010	1	0	1	0	0	0	0	0	1	0	1							3
1011	1	1	1	1	0	0	0	0	1	1	1	1						7
1100	1	0	0	0	1	0	0	0	1	0	0	0	1					3
1101	1	1	0	0	1	1	0	0	1	1	0	0	1	1				7
1110	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1			7
1111	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		15
10000	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1

Therefore, as shown in Ref. 1,

$$\begin{aligned} b_k &= h(a'_{k-1}, a'_{k-2}, \cdots, a'_{k-r-1}) + a'_{k-r} \\ &= h(a_{k-1}, a_{k-2}, \cdots, a_{k-r-1}) + a_{k-r} + 1 \\ &= a_k + 1 = a'_k \end{aligned}$$

If the states $a_{k-1}, a_{k-2}, \cdots, a_{k-r}$ and $a'_{k-1}, a'_{k-2}, \cdots, a'_{k-r}$ lie on different cycles and have the same length, their cycles are complementary images of one another. Assuming the FSR's with an $f_1(x)$ of Eq. (4) yield cycles of equal length, complementary images of each cycle will lie on another cycle when r does not divide s (see Eq. 5). When r equals 2^q (and thus divides s) complementary states will lie in the same cycle and will be separated by 180 deg.

d. Cycle length versus initial state. The cycle length of the longest cycle for $f_1(x)$ of Eq. (4) was determined in Subsection 2-b. It remains to show that none of the initial states result in a $g_1(x)$ that has a factor common with $f_1(x)$. These states then must all lie in cycles of length equal to the one containing the $00 \cdots 0$ state.

Let

$$r = 2^q - 1$$

then

$$(x+1)^r = \sum_{i=0}^{r} x^i \tag{6}$$

Equation (6) follows from

$$\frac{(x+1)^{2^q}}{x+1} = \frac{x^{2^q}+1}{x+1}$$

The numerator of $G_0(x)$ in (2) gives $g_0(x)$ as follows:

$$\left.\begin{array}{l} a_{-1} \\ + a_{-2} + a_{-1}x \\ + \quad \cdot \\ \quad\ \cdot \\ \quad\ \cdot \\ + \ a_{-2^q+1} + a_{-2^q+2}\,x + \cdots + a_{-1}\,x^{2^q-2} \end{array}\right\} \tag{7}$$

The mod 2 summation of all the terms in Eq. (7) is $g_0(x)$. The expression for $g_1(x)$ is

$$1 + (x+1)\,g_0(x)$$

The tabulation of terms comprising $g_1(x)$ is

$$\left.\begin{array}{l} 1 \\ + a_{-1} \\ + a_{-2} \\ \quad\ \cdot \\ \quad\ \cdot \\ \quad\ \cdot \\ + a_{-2^q+1} + a_{-2^q+1}\,x + \cdots + a_{-2}\,x^{2^q-2} + a_{-1}\,x^{2^q-2} \end{array}\right\} \tag{8}$$

For $g_1(x)$ to have a common factor with $f_1(x)$, it must be of the form $(x+1)^t$ where $t < r+1$. Recall that $(x+1)^t$ results in a polynomial with a constant term 1 and an odd number of terms which are distinct powers of x. The latter condition can be satisfied by assigning an odd number of a_{-i}'s the state value 1. This leads to a contradiction, since the constant term

$$1 = 1 + a_{-1} + a_{-2} + \cdots + a_{-2^q+1}$$

and requires that an even number of a_{-i}'s be assigned the state value 1. Therefore, there is no initial state which results in a $g_1(x)$ that has a common factor with $f_1(x)$. The same argument can be applied for other values of r.

Example 3. Given

$$\begin{aligned} f_1(x) &= (x+1)(x+1)^5 \\ &= (x-1)(x^5 + x^4 + x + 1) \end{aligned}$$

the terms comprising $g_0(x)$ are tabulated as follows:

$$\begin{array}{l} a_{-1} \\ a_{-4} + a_{-3}\,x + a_{-2}\,x^2 + a_{-1}\,x^3 \\ a_{-5} + a_{-4}\,x + a_{-3}\,x^2 + a_{-2}\,x^3 + a_{-1}\,x^4 \end{array}$$

The terms comprising $g_1(x)$ are

$$\begin{array}{l} 1 + a_{-1} + a_{-4} + a_{-5} \\ + (a_{-1} + a_{-2} + a_{-5})\,x \\ + (a_{-2} + a_{-4})\,x^2 \\ + (a_{-1} + a_{-3})\,x^3 \\ + a_{-2}\,x^4 \\ + a_{-1}\,x^5 \end{array}$$

To reduce $f_1(x)$ to $x + 1$, $(x + 1)^2$, or $(x + 1)^3$, $g_1(x)$ would have to be $(x + 1)^5$, $(x + 1)^4$, or $(x + 1)^3$, respectively. For $g_1(x)$ of

(1) $(x + 1)^5 = x^5 + x^4 + x + 1$,

$$a_{-1} = a_{-2} = a_{-3} = a_{-4} = a_{-5} = 1$$

but

$$1 \neq 1 + a_{-1} + a_{-4} + a_{-5}$$

(2) $(x + 1)^4 = x^4 + 1$,

$$a_{-2} = a_{-4} = 1$$

$$a_{-1} = a_{-3} = a_{-5} = 0$$

but

$$1 \neq 1 + a_{-1} + a_{-4} + a_{-5}$$

(3) $(x + 1)^3 = x^3 + x^2 + x + 1$

$$a_{-3} = a_{-4} = 1$$

$$a_{-1} = a_{-2} = a_{-5} = 0$$

but

$$1 \neq 1 + a_{-1} + a_{-4} + a_{-5}$$

The foregoing procedure for determining whether or not an initial state results in a $g_1(x)$ which has a common factor with $f_1(x)$ is prohibitive when r does not equal $2^q - 1$.

The foregoing proof for any r equal $2^q - 1$ is a starting point for showing that cycles associated with r of $2^q - 2$ are equal in length, and so on. The following relationships hold:

$$\{a_k^r\} + \{a_{k-1}^r\} = \{a_k^{r-1}\}$$

where the superscript r denotes the number of stages in the FSR, with an $f_1(x)$ equal to $(x + 1)(x + 1)^r$.

Also

$$\{1 + a_k^r\} + \{1 + a_{k-1}^r\} = \{a_k^{r-1}\}$$

$$\{a_k^r\} + \{a_{k-2^r}^r\} = \{a_k^{r-2^r}\}$$

(See Table 3.)

Table 3. Equal-length cycles for r of 2, 3, 4, and 5

1 2 3 4 5	a_k^5	1 2 3 4	a_k^4	1 2 3	a_k^3	1 2	a_k^2
0 0 0 0 0	1	0 0 0 0	1	0 0 0	1	0 0	1
1 0 0 0 0	0	1 0 0 0	1	1 0 0	0	1 0	1
0 1 0 0 0	1	1 1 0 0	1	0 1 0	0	1 1	0
1 0 1 0 0	0	1 1 1 0	1	0 0 1	0	0 1	0
0 1 0 1 0	0	1 1 1 1	0				
0 0 1 0 1	0	0 1 1 1	0	1 1 1	0		
0 0 0 1 0	0	0 0 1 1	0	0 1 1	1		
0 0 0 0 1	0	0 0 0 1	0	1 0 1	1		
				1 1 0	1		
1 1 1 1 1	0						
0 1 1 1 1	1						
1 0 1 1 1	0						
0 1 0 1 1	1						
1 0 1 0 1	1						
1 1 0 1 0	1						
1 1 1 0 1	1						
1 1 1 1 0	1						
0 0 1 1 0	0	0 1 0 1	0				
0 0 0 1 1	1	0 0 1 0	1				
1 0 0 0 1	1	1 0 0 1	0				
1 1 0 0 0	0	0 1 0 0	1				
0 1 1 0 0	1	1 0 1 0	1				
1 0 1 1 0	1	1 1 0 1	0				
1 1 0 1 1	0	0 1 1 0	1				
0 1 1 0 1	0	1 0 1 1	0				
1 1 0 0 1	1						
1 1 1 0 0	0						
0 1 1 1 0	0						
0 0 1 1 1	1						
1 0 0 1 1	0						
0 1 0 0 1	0						
0 0 1 0 0	1						
1 0 0 1 0	1						

Cycles of equal length are shown on Good's diagram for an r of 1, 2, 3, and 4 (Fig. 2). Each vertex of a Good diagram (Ref. 3) represents a state of an FSR. Two edges enter and two edges leave every vertex. An edge leaves a vertex and enters a vertex which is a possible successor. Each vertex thus has two possible successors and two possible predecessors.

References

1. Golomb, S. W., *Shift Register Sequences*. Holden-Day, Inc., San Francisco, Calif., 1967.

References (contd)

2. Hardy, G. H., and Wright, E. M., *An Introduction to the Theory of Numbers.* Oxford University Press, Amen House, London, 1938.

3. Good, I. J., "Normal Recurring Decimals," *J. London Math. Soc.*, Vol. 21, Part 3, pp. 167–169, 1946.

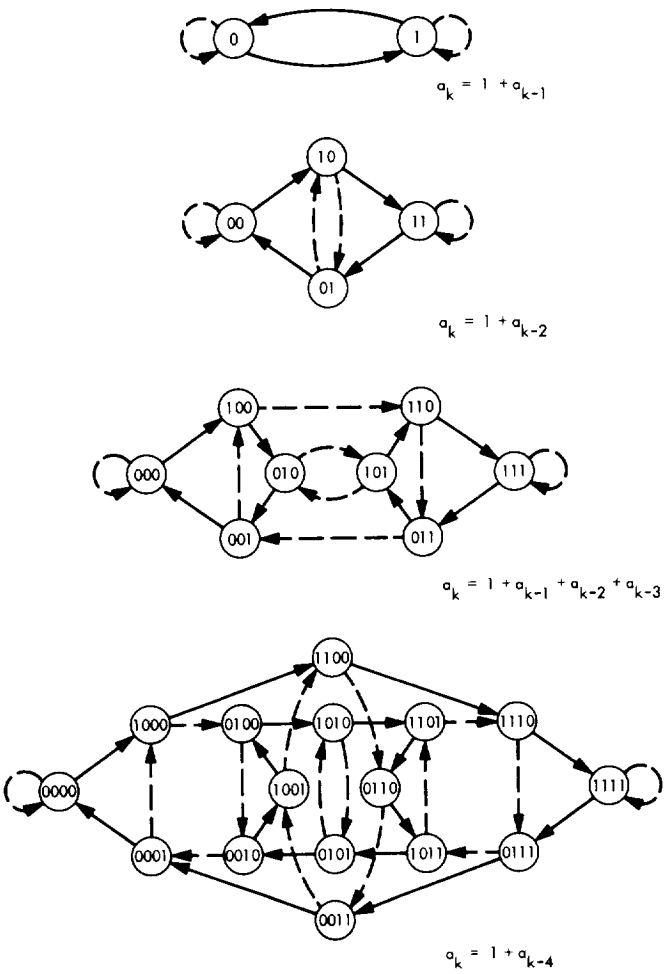


Fig. 2. Good's diagram for r of 1, 2, 3, and 4

B. Capsule System Advanced Development Entry Data Subsystem, *R. V. Gutierrez*

1. Introduction

The entry data subsystem (EDS)[1], a subsystem within the entry capsule system, has been designed and developed in accordance with the requirements set forth by the Capsule System Advanced Development (CSAD) program.

[1]Casani, K., *Capsule System Advanced Development Program Report,* July 15, 1968 (JPL internal document).

The EDS is designed to accept digital and analog information from all the entry capsule's science and engineering sources, process it, format it, and send it to the spacecraft for transmission to earth.

All analog inputs are connected to a 62-channel multiplexer for subsequent analog-to-digital conversion, using an analog-to-pulse width converter (A/PW) (Ref. 1) and conversion scheme. This data is fed to the spacecraft prior to separation, to the entry radio subsystem for transmission back to the spacecraft after separation and during entry, and to the EDS's plated-wire memory-storage unit during entry. Storing data in the plated-wire memory-storage unit assures that the accelerometer, radiometer, and engineering data obtained during the radio blackout period of entry is not lost and will be retrieved after blackout and prior to impact.

The EDS is designed to interface with the following scientific and engineering data sources:

(1) The JPL mass spectrometer designed to perform a compositional analysis of the atmospheric envelope of the planet Mars.

(2) The water vapor detection system designed to investigate the water content of the Mars atmosphere during capsule entry.

(3) The aerometry experiment package (consisting of accelerometers, pressure and temperature detectors) designed to determine the gas density, pressure, and temperature as a function of altitude above the surface of the planet Mars.

(4) The radiometer experiment designed to determine the principal chemical constituents of the Mars atmosphere and the mole fraction of these constituents from radiometric measurements of the intensities of selected bands and lines in the emission spectrum of the high-temperature shock layer formed ahead of the probe during entry.

(5) All the analog transducers from the entry capsule's 62 science and engineering sources.

There was a requirement for a storage unit of the correct size to accommodate the data from the maximum calculated blackout period.

The assumption was made that the Mach 9 signal would be a true indication that the blackout period is

over. Therefore the EDS will start writing into memory at E−12 (12 min before entry) and start playing back from memory at Mach 9, after calibration is over.

A breadboard EDS using Signetic integrated circuits and operational support equipment (OSE) have been designed and fabricated at JPL. They were tested and delivered to the Spacecraft Assembly Facility (SAF) for integration within the CSAD entry system configuration on March 21, 1968.

2. Functional Design

a. Operation. The EDS will always be in the entry mode of operation at power turn-on. Figure 3 is a time-phase profile for the two modes of operation.

The EDS will be powered on at S−175 (175 min before separation) and operate in accordance with the CSAD entry format (Fig. 4). At S−120 (120 min before separation) the EDS will receive a command from the entry sequencer and timer subsystem (ES&T) and switch into the separation mode of operation. In this mode, the EDS will operate in accordance with the CSAD separation format and sample only engineering sensors. The EDS will be powered off at S+15 (15 min after separation).

The EDS will be powered on at E−15 (15 min before entry) and operate in accordance with the entry format. At E−14, the EDS issues calibrate commands to the aerometry experiment, the radiometer and the water vapor experiment. At E−12 the EDS starts writing into memory in accordance with the EDS *write blackout* format (Fig. 5). Upon receiving a *Mach 9* command, the EDS stops writing into memory, issues a *radiometer calibrate* command and starts reading out from memory. In this mode the EDS operates in accordance with the entry format (Fig. 4) with the memory information substituted in the radiometer slots.

b. Modes of operation. The EDS has essentially two modes of operation: the separation mode and the entry mode. Writing into memory and reading from memory occur only in the entry mode of operation.

Separation mode. The EDS receives a separation command from ES&T and switches into the separation mode of operation. In this mode the EDS will operate in accordance with the separation format, and will sample only engineering sensors. All measurements are 6-bit measurements, except for pseudonoise (PN) (15 bits), frame count (9 bits), and Accutron clock timer (10 bits).

The 3-sec frame is divided into 1-sec subframes, and each subframe is divided into 30 units. Each unit is, in turn, divided into three 6-bit measurements; thus, the total number of bits for a 1-sec subframe is

$$3 \times 6 \text{ bits} \times 30 \text{ units} = 540 \text{ bits/sec}$$

All subframes are identical, except for the first 24 bits of subframe 1.

Entry mode. In this mode the EDS operates in accordance with the entry format of Fig. 4. This format is divided into three essentially identical subframes. The radiometer, temperature, and pressure are 7-bit measurements; the water vapor detector is a 9-bit measurement; the accelerometers are 10-bit measurements; and the mass spectrometer is a 22-bit measurement.

The EDS *write blackout* format is a 1-sec, 224-bit frame and is shown in relationship with the CSAD entry 1-sec subframe in Fig. 5.

Upon receiving a *Mach 9* command from the parachute initiator subsystem, the EDS stops writing into memory, issues a *radiometer calibrate* command, and starts to read out the blackout information from memory at 112 bits/sec, according to the CSAD entry format of Fig. 4. Blackout memory data is then substituted for the radiometer data within the radiometer time slots of the entry format. It should be noted that blackout information is read out in the following manner: last in, first out.

c. System block diagram. The block diagram of the complete EDS is shown in Fig. 6. The EDS may be divided into eight functional blocks:

Oscillator and timing generator. The EDS timing generator contains a 1.728-MHz crystal oscillator whose output feeds a countdown timing chain. Outputs from this countdown chain are decoded to provide all of the timing requirements of the EDS.

Instrument control. The instrument control receives all the required timing signals and generates commands that control all the instruments that interface with the EDS.

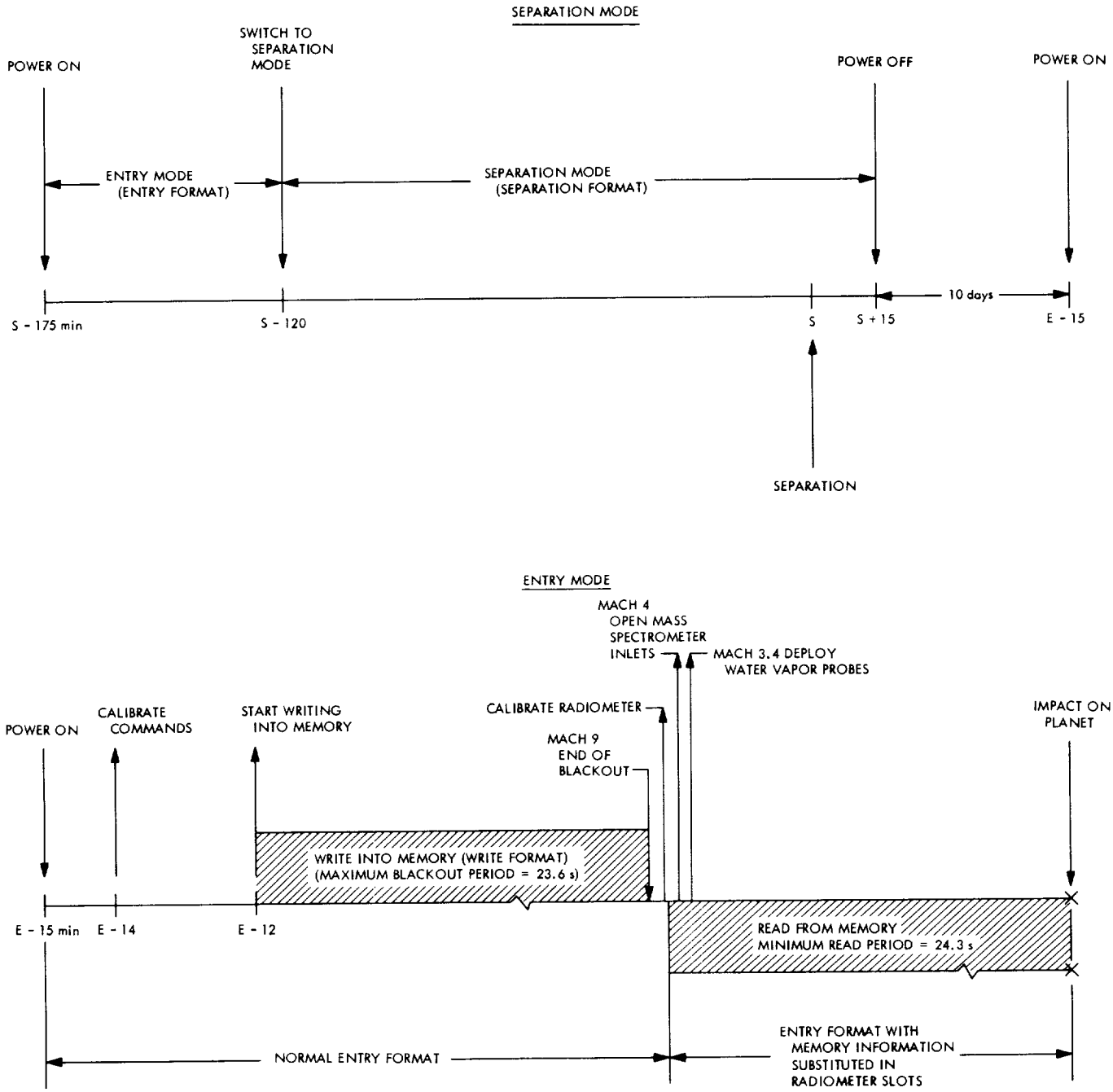


Fig. 3. Time-phase profile of EDS

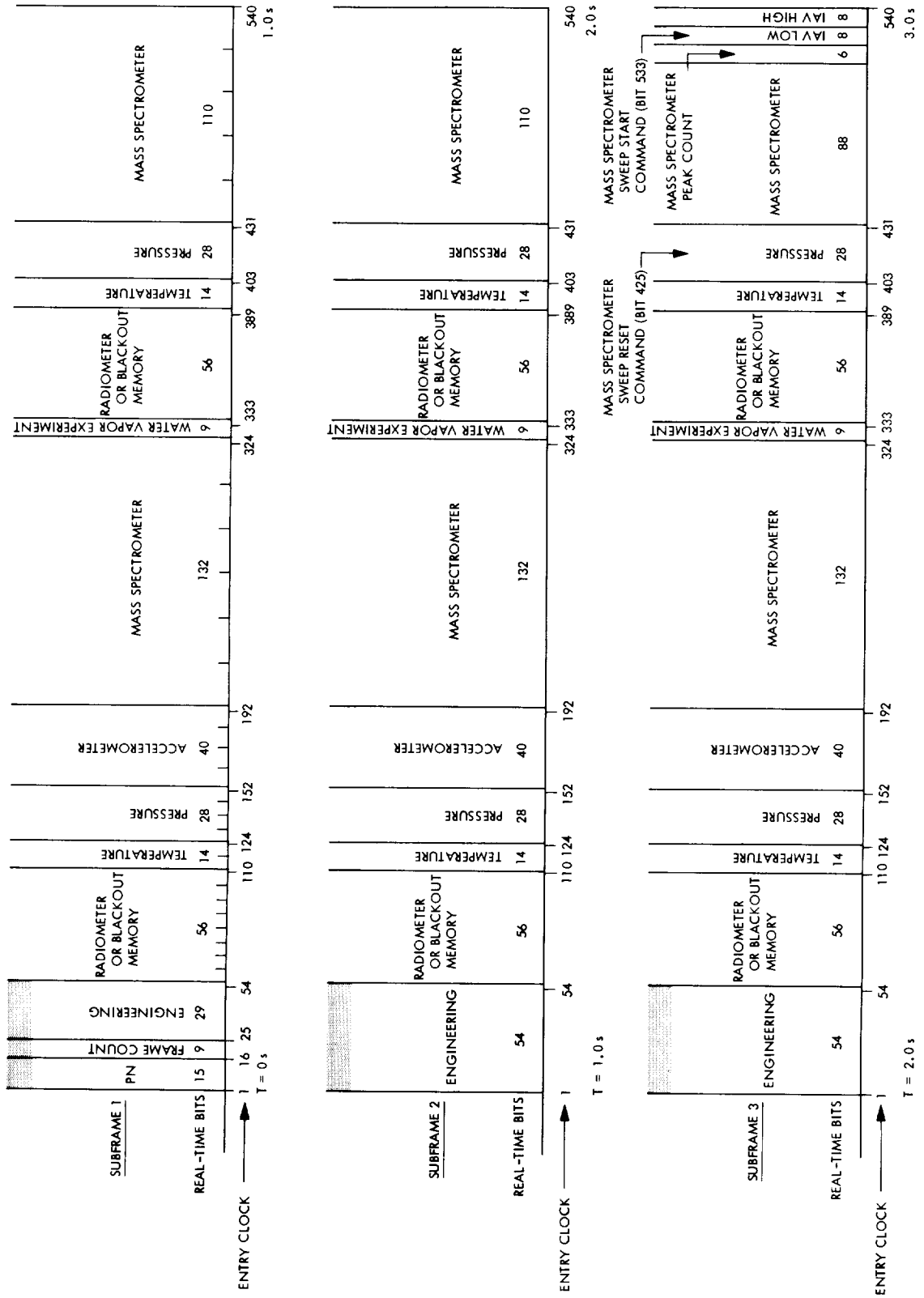


Fig. 4. CSAD entry format

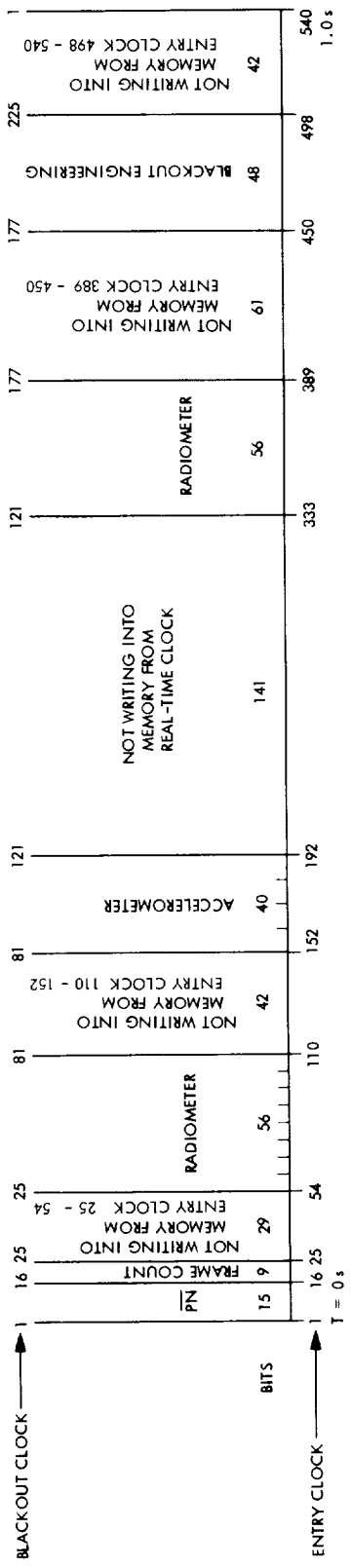

BLACKOUT CLOCK
ENTRY CLOCK
BITS
T = 0 s
1.0 s
NOT WRITING INTO MEMORY FROM ENTRY CLOCK 498 - 540
BLACKOUT ENGINEERING
NOT WRITING INTO MEMORY FROM ENTRY CLOCK 389 - 450
RADIOMETER
NOT WRITING INTO MEMORY FROM REAL-TIME CLOCK
ACCELEROMETER
NOT WRITING INTO MEMORY FROM ENTRY CLOCK 110 - 152
RADIOMETER
NOT WRITING INTO MEMORY FROM ENTRY CLOCK 25 - 54
FRAME COUNT
PN

Fig. 5. EDS *write blackout* format

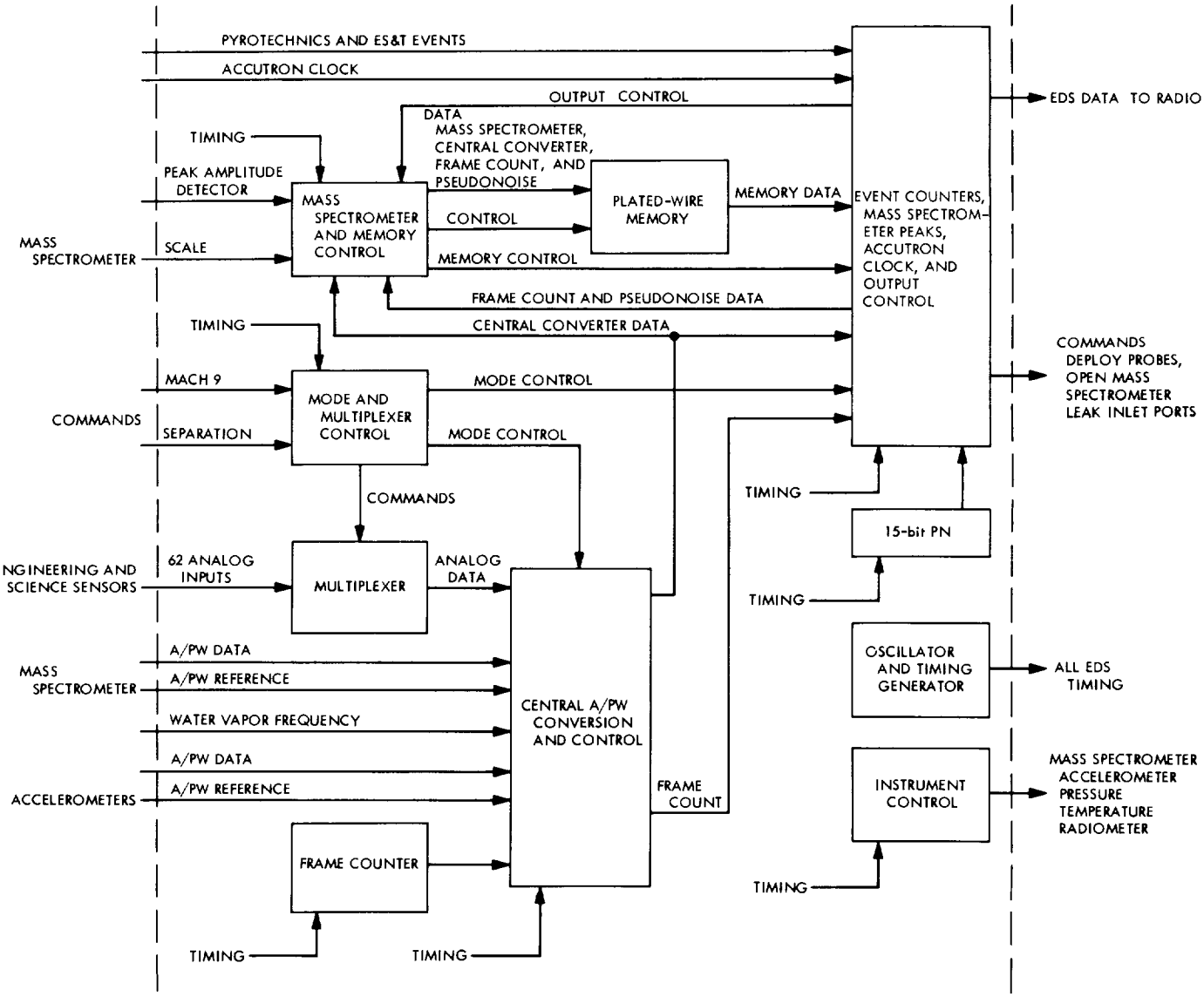


Fig. 6. EDS block diagram

Multiplexer. All the 62 analog inputs from the engineering and science sensors are connected to the EDS multiplexer. Upon command from the multiplexer control, these analog inputs are normalized to 6 V, then switched to the central analog-to-pulsewidth (A/PW) converter for analog-to-digital conversion. The timing requirements for the multiplexer are dictated by the mode of operation of the EDS and the formats required by these modes. The detailed design and description of the EDS multiplexer is described in SPS 37-50, Vol. III, p. 216.

Multiplexer and mode control. The multiplexer control logic consists of timing signals: entry, blackout, and separation—three unique gate matrices which feed the FET control matrix and the output control logic. The FET control matrix is a gate matrix which receives the timing signals and issues all the commands required by the multiplexer.

The mode control logic consists of gates and flipflops which set the mode of operation of the EDS. The EDS receives only two external commands: the *separation* command and the *Mach 9* indicator command. The separation command causes the EDS to switch from the entry mode to the separation mode of operation. The *Mach 9* command allows the Mach 9 counter to count, the *calibrate* commands to be sent, the memory to start

reading back, the *probe deploy* and *open inlet port* commands to be sent.

Central A/PW conversion and control. The analog-to-digital conversion is achieved within the EDS by using a *Mariner* flight-qualified A/PW converter. The A/PW converter converts analog measurements to digital form with a resolution of 1 part in 511 (9 bits).

The output pulses of the A/PW (data and reference), the mass spectrometer A/PW pulses, the accelerometer A/PW pulses, and the water vapor frequency input are fed into the A/PW control logic with the appropriate timing signals. The control logic allows the high-frequency (864 kHz) clock to be counted by the central counter for a period of time equivalent to the delay between the A/PW reference and data pulses. This time period is directly proportional to the analog input voltage at the A/PW converter (Ref. 1). For the water vapor experiment, the control logic allows the water vapor frequency to be counted for a fixed period of time. The result of this count is the central converter data and is sent to the memory control and the output control units in accordance with the mode and formats in use. To ensure correct operation, the logic, counters, and shift registers are kept in sync by being reset at the start of an A/PW conversion period or at the start of a water-vapor reading period.

Memory and mass spectrometer control. The memory control logic receives central converter data, mass spectrometer data, frame count and PN data, as well as appropriate timing signals. Its primary function is to condition this incoming data for subsequent storage within the plated-wire memory. It also provides the memory with the read/write control function, the clock, and the address location lines.

The mass spectrometer control logic receives the peak detect signal and the scale inputs from the instrument. Its primary function is to convert this peak signal (11-bit conversion), using a similar type of conversion scheme to that used for A/PW conversions. It also provides a time of occurrence measurement (9 bits) indicating when this peak occurs with respect to the sweep start of the instrument, and scale measurements (2 bits) indicating the range of the peak measurement.

Each time a peak detect signal is received, the mass spectrometer control logic outputs 22 bits of information for storage within the plated-wire memory and counts the number of peaks received during the given sweep time of the instrument.

Plated wire memory. This unit was designed and developed by Librascope under JPL contract with in-house JPL design support. It receives data, a read/write control signal, a clock, and 13 address lines from the memory control unit. Depending upon the status of the control signal and the receipt of a clock, the memory unit will either read information from an address location or write data into a location within the stack. For a detailed description of the CSAD woven plated-wire memory unit, refer to SPS 37-51, Vol. III, p. 175.

Event counters and output control. There are four event counters which monitor pyrotechnics and ES&T events. As an event occurs, these counters are updated by one, and at the appropriate time and in accordance with the mode and data format in question, the content of the counters are transferred to the output control logic.

The output control logic receives ES&T clock counter data, mass spectrometer peak counter data, pyrotechnics and ES&T data, memory data, central converter data, mode data, frame count and PN data. With the appropriate timing, it outputs this data in accordance with the mode of operation and the format under consideration. All EDS data (except PN) is fed into a half adder. The relay receiver bit synchronizer on the parent spacecraft will not operate correctly when it receives a long string of zeros. Due to the fact that the mass spectrometer data may be a long series of over 100 *zeros* (no peaks detected), the requirements of half-adding *ones* and *zeros* to the EDS data was imposed. The outputs of the EDS half-added data and the output from the 15-bit PN generator are *or*-gated together and sent to the radio subsystem. This resultant EDS data is used as the modulating signal in the radio subsystem.

3. Testing

The CSAD EDS underwent two levels of room temperature testing.

a. Initial checkout. The EDS underwent initial checkout with its operational support equipment (OSE). The OSE contains all the power supplies required by the EDS, the memory unit, and the OSE. It also contains a modified 40 line/sec Franklin printer which, in conjunction with its logic, is capable of printing out in real-time all the data that the EDS sends to the entry radio subsystem or all the data that the EDS OSE receives from the relay receiver demodulation subsystem.

The OSE logic is designed to simulate all system inputs to the EDS (Fig. 6). Switches on the front panel

control the number of mass spectrometer peaks, and the width of these peaks sent to the EDS. Pushbuttons simulate the events that are counted, and the *Mach 9* and *separation* commands. The A/PW data and reference pulses are simulated and controlled by thumbwheels on the OSE front panel. All 62 analog inputs are simulated and fed into the EDS; voltage potentiometers control the analog inputs through a known voltage range. Lamps on the front panel indicate the status of the EDS as it responds to a command, changes in mode of operation, etc.

b. System functional testing. Upon delivery to the Spacecraft Assembly Facilities (SAF), the EDS and the other subsystems of the CSAD Capsule System underwent capsule system functional testing.

c. Test results. After the initial debugging during checkout, the EDS operated correctly with all the simulated inputs from the OSE.

Throughout SAF testing the EDS operated correctly, without any problems. The data was good, with occasional noise glitches that caused errors in the data retrieval system contained within the EDS's OSE system. Upon further investigation it was determined that these glitches occurred when the mass spectrometer ion pump was turned on. It was also determined that they were the result of a combination of the system configuration at SAF, the noise sensitivity of the data retrieval system within the EDS's OSE, and the noise effect on the system upon turning on the ion pump.

4. Future Activity

A general review of the EDS logic should be conducted to ensure simplification in design, minimization of components and power consumption, and maximization of reliability.

A review of the memory and its electronics is in order so as to improve the margins and noise immunity of its input circuits. The memory should also be designed mechanically to withstand shock and vibration as required in a flight project.

The EDS multiplexer presently uses PF 157 flatpack which contains four 2N3386 p-channel junction field-effect transistors (JFET). These JFET's are employed as switches, but unfortunately they are normally closed (when power is off), thus shorting all the analog inputs through high-impedance paths. This situation may not be acceptable in a flight configuration, and new designs for the multiplexer switches are needed.

Reference

1. Nixon, R. H., *Bipolar Analog-to-Pulse Width Converter,* Technical Report 32-1034. Jet Propulsion Laboratory, Pasadena, Calif., Jan. 15, 1967.

XV. Lunar and Planetary Instruments and Sciences

SPACE SCIENCES DIVISION

A. Infrared Absorption Spectrum of CH_4 at 9050 cm^{-1}, *J. S. Margolis and K. Fox*[1]

A high-resolution infrared absorption spectrum has been obtained for the region of the R branch of the second overtone of the triply degenerate fundamental ν_3 of CH_4. The measured lines between 9057 and 9133 cm^{-1} are reproduced in Fig. 1. Comparisons have been made between the tetrahedral fine structure splittings in this spectrum and those in the spectra of ν_3 and $2\nu_3$ of CH_4 and $2\nu_3$ of CD_4. A more detailed account will appear in *The Journal of Chemical Physics* (September 1, 1968).

[1]NRC–NASA resident research associate on leave from The University of Tennessee.

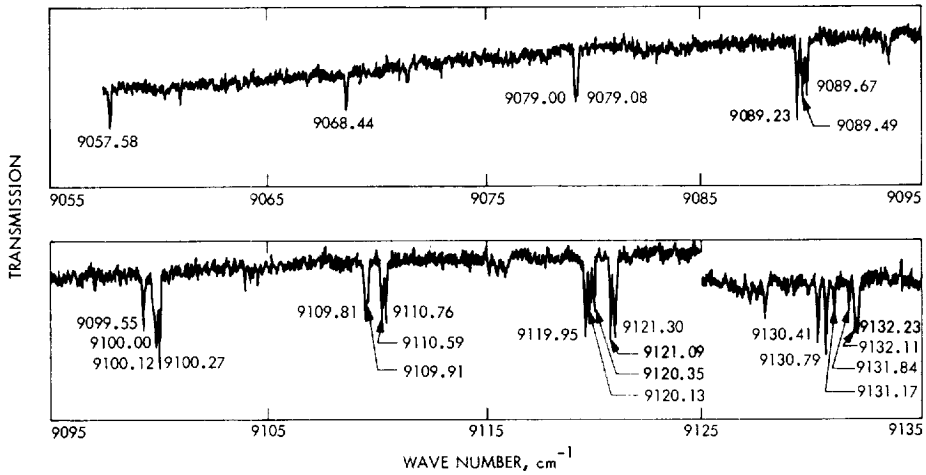


Fig. 1. Absorption spectrum of CH_4 in R branch of 9050 cm^{-1} band (75 torr, 64 m, 294°K)

B. Peak Detector–Analog/Pulse Width Converter Analysis, *J. R. Locke*

1. Introduction

The peak detector–analog/pulse width converter (PD–A/PWC) has potential application in spaceborne gas chromatography and mass spectrometry instrumentation. It was developed by JPL in discrete component form and has been successfully interfaced with a mass spectrometer to provide signal conditioning in a form appropriate for digital processing that simultaneously furnishes qualitative (peak time of occurrence) and quantitative (pulse width proportional to signal voltage peak) information. It has been operated over the temperature range of 0 to 50°C, detecting signals ranging from 10 mV to 10 V. It has good linearity, between pulse width and peak input, deviating over the range of 0.1 to 10 V less than 0.15%. A novel feature of the circuit is its noise immunity, which is a minimum of 17 mV and can be increased by making the static offset larger. Worst-case peak-detection error occurs for low-level triangular input signals and is less than 3 ms.

To maximize the value of this circuit, it is planned that the final configuration will be in microminiature form. In anticipation of making this conversion, the Planning Research Corporation (PRC) of Los Angeles, California was contracted to perform a computer-aided analysis of the circuit that would provide analytical guidelines for the microminiaturization. The decision to utilize computer simulation methods was motivated by the following considerations:

(1) They provide a thorough means of investigating circuit performance that would be prohibitively time-consuming if sought through standard breadboarding techniques.

(2) They are much faster and more accurate than hand analysis.

(3) They allow (equivalent) measurements that would be extremely difficult to make in an actual circuit (such as a low-level voltage measurement on a high-impedance point, or the values of all node voltages at the peak of a signal).

Two computer programs were used in the analysis:

(1) PRONE—a nonlinear dc circuit analysis program developed by PRC.

(2) SCEPTRE—a transient analysis program developed by the U.S. Air Force Weapons Laboratories.

It should be noted that complementary breadboard testing was used in conjunction with the computer simulations to confirm their validity.

The material covered in this article is an abstract of the final report prepared by PRC. The complete report will be published as a technical memorandum in October 1968.

2. Functional Description of the PD–A/PWC

Figure 2 is a complete schematic of the PD–A/PWC. Peak detection is achieved by storing a voltage proportional to the peak on capacitor C; as the input voltage drops below the peak, CR1 is disconnected creating an unbalance across matched transistors Q1AB which toggles Q5 from a normally high to a low state. When the output goes low, it provides a signal that fires a *one* shot. The *one* shot controls a series-shunt switch that disconnects the input from the signal source for a period of time long enough to allow discharge of the maximum input signal seen by capacitor C. By discharging C with a constant current source Q6, until Q5 reverts to its original state, the resultant pulse width defined by Q5 is proportional to the peak input voltage.

3. Analysis Objectives

The main objectives of the analysis were to determine what parameters were of primary importance in controlling:

(1) Peak detection.

(2) Peak-to-pulse-width conversion accuracy.

Four areas were explored in seeking this information:

(1) The leading edge trigger.

(2) The capacitor discharge circuit.

(3) The trailing edge trigger.

(4) The interrelationship of these factors as they affect peak detection and pulse width.

Another area that was given limited treatment was input noise sensitivity.

Preliminary hand analysis showed the effects of the series-shunt switch and the *one* shot to be of a secondary significance. It was also observed that the buffer amplifier could affect circuit performance in terms of its input and

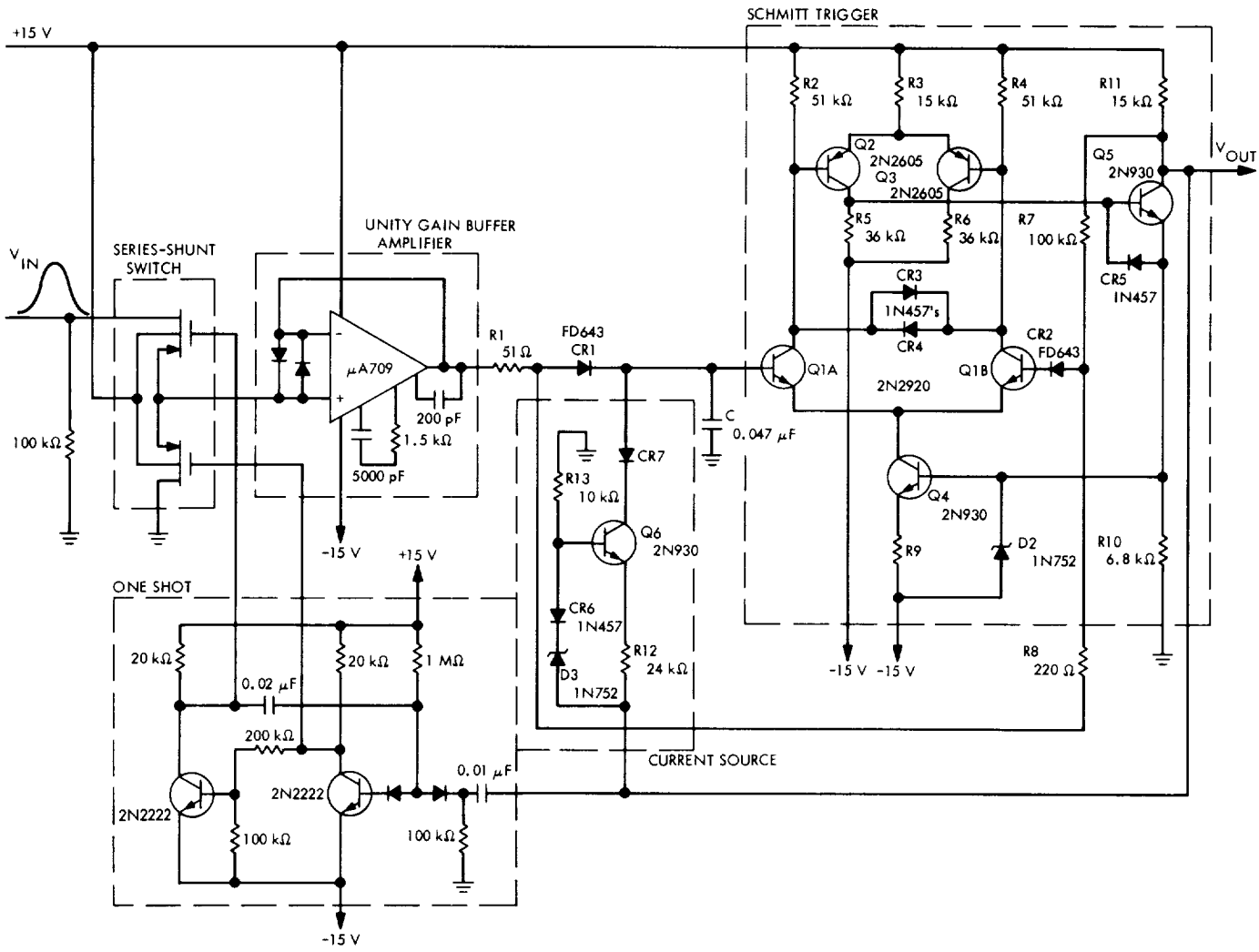


Fig. 2. Complete peak detector–analog/pulse width converter

output impedance and input offset voltage, but that these factors could be handled in a straightforward manner with standard analysis techniques. For these reasons, the series-shunt switch, *one* shot, and buffer amplifier were not treated in the computer simulations.

4. Circuit Analysis

a. Leading edge trigger. The leading edge trigger analysis is concerned with two considerations associated with the instant when the input signal produces a change of state at the output:

(1) The Schmitt trigger changes state at some non-zero base voltage differential. This voltage is a function of CR1, CR2, Q1AB, R1, R7, R8, R11, positive voltage V_{POS}, input voltage V_{IN}, temperature, and to some degree the remaining components in the circuit.

(2) The rate of change of the input signal will cause a component of current to flow in CR1 that is not present in CR2. In addition to making the drops across CR1 and CR2 unequal, the rate-of-change current component affects the dynamic resistance of CR1 and thus the manner in which capacitor C charges.

To study the first of these two considerations, the Schmitt trigger portion of the PD–A/PWC was simulated together with a fictitious negative feedback circuit. The feedback circuit senses Q5 threshold current and feeds back a voltage V_f' corresponding to that which capacitor C would have to attain to cause toggling for various

input voltage V^1_{IN} values.[2] A second simulation of the Schmitt trigger circuit was run without the negative feedback circuit, to determine the peak input voltages V_P corresponding to the capacitor trigger voltages V^1_C. The difference between the peak input V_P voltage values and their corresponding trigger input voltages V^1_{IN} is the static offset voltage error V^S_{OS}, which is the minimum voltage difference required to produce toggling. Such a static offset error would occur if the input signal flat-topped and remained at its peak value long enough for capacitor C to reach its steady-state value.

Figure 3 illustrates the meaning of these terms, and Table 1 summarizes the computed data for the circuit at room temperature. A plot of V^1_{IN} versus V^1_C results in the trigger locus, which will be combined later with the transient trajectory to define the PD–A/PWC's dynamic behavior. Table 2 summarizes the PD–A/PWC's static offset voltage sensitivity to parameter variation. The parameters considered are resistors, voltage sources, base-emitter junctions, junction leakages, and transistor current gains h_{fe} at peak input voltages V_P of 0 and 10 V.

[2]Numerical superscripts are defined as follows:

X^1 = quantity obtained in first computer simulation to analyze the leading edge static trigger.

X^2 = quantity obtained in second computer simulation to analyze the Q5 *off* equivalent circuit.

X^3 = quantity obtained in third computer simulation to analyze the capacitor charging circuit.

X^4 = quantity obtained in fourth computer simulation to analyze the trailing edge trigger.

X^5 = quantity obtained in fifth computer simulation to analyze the capacitor discharge circuit.

X^6 = quantity obtained in sixth computer simulation to analyze Q5 *on* equivalent circuit.

Table 1. Nominal 25°C static offset voltage and trigger locus data

V_P, V[a]	V^1_{IN}, V[b]	V^S_{OS}, V[c]	$V^2_C(V_P)$, V[d]
0.0	−0.0167290	0.0167290	−0.2635916
0.5	0.4838834	0.0161166	0.2358232
1.0	0.9844980	0.0155020	0.7352401
1.5	1.4851144	0.0148856	1.2346588
2.0	1.9857331	0.0142669	1.7340798
2.5	2.4863538	0.0136462	2.2335027
3.0	2.9869765	0.0130235	2.7329277
3.5	3.4876014	0.0123986	3.2323549
4.0	3.9882282	0.0117718	3.7317840
4.5	4.4888572	0.0111428	4.2312157
5.0	4.9894880	0.0105120	4.7306483
5.5	5.4901214	0.0098786	5.2300839
6.0	5.9907564	0.0092436	5.7295212
6.5	6.4913939	0.0086061	6.2289609
7.0	6.9920331	0.0079669	6.7284024
7.5	7.4926745	0.0073255	7.2278460
8.0	7.9933180	0.0066820	7.7272917
8.5	8.4939635	0.0060365	8.2267395
9.0	8.9946112	0.0053888	8.7261894
9.5	9.4952610	0.0047390	9.2256414
10.0	9.9959130	0.0040870	9.7250959

[a]V_P is the peak value of the input voltage.

[b]V^1_{IN} is the voltage the input must drop to before triggering will occur.

[c]$V^S_{OS} = V_P - V^1_{IN}$ when V_{IN} has sat at V_P long enough for capacitor C to attain its steady-state value.

[d]$V^2_C(V_P)$ is the steady-state voltage on capacitor C.

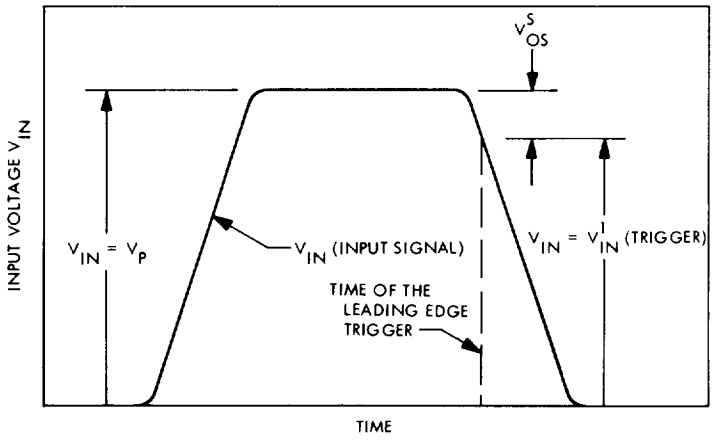


Fig. 3. Input signal for defining static offset terminology

Table 2. Static offset sensitivities

Parameter	Peak input voltage V_P			
	$V_P = 0$ V		$V_P = 10$ V	
Resistor sensitivity S_R				
Resistor	S_R, V/ %	Rank	S_R, V/ %	Rank
R1	0.000×10^{-1}	–	-0.001×10^{-4}	–
R2	-0.490×10^{-3}	1	-0.510×10^{-3}	1
R3	0.149×10^{-4}	–	0.149×10^{-4}	–
R4	0.472×10^{-3}	2	0.472×10^{-3}	2
R5	-0.163×10^{-4}	–	-0.163×10^{-4}	–
R6	-0.431×10^{-10}	–	-0.431×10^{-10}	–
R7	-0.140×10^{-3}	4	-0.345×10^{-4}	4
R8	0.159×10^{-3}	3	0.388×10^{-4}	3
R9	0.190×10^{-4}	–	0.019×10^{-3}	–
R10	0.000	–	0.000	–
R11	0.368×10^{-4}	5	-0.265×10^{-4}	5
RL	0.182×10^{-4}	–	0.216×10^{-4}	–
Emitter junction sensitivity S_E				
Transistor	S_E, mV/mV	Rank	S_E, mV/mV	Rank
Q1A	0.62835479	2	0.67113600	2
Q1B	−0.62835594	1	−0.67115959	1
Q2	0.01010435	4	0.01010583	3
Q3	−0.00974759	5	−0.00974899	4
Q4	0.00046383	6	0.0004729	6
Q5	−0.01781465	3	−0.0041525	5
Diode junction sensitivity S_D				
Diode	S_D, mV/mV	Rank	S_D, mV/mV	Rank
CR1	0.628	1	0.671	1
CR2	−0.628	1	−0.699	2
CR3	0.195×10^{-4}	3	0.195×10^{-4}	4
CR4	-0.259×10^{-4}	2	-0.259×10^{-4}	3
CR5	0.241×10^{-9}	4	0.221×10^{-9}	5

Parameter	Peak input voltage V_P			
	$V_P = 0$ V		$V_P = 10$ V	
Junction leakage sensitivity S_L				
Diode/ transistor	S_L, μV/μA	Rank	S_L, μV/μA	Rank
CR3	1.01×10^{3}	3	1.01×10^{3}	3
CR4	-1.01×10^{3}	3	-1.01×10^{3}	3
Q1AC	-71.5×10^{3}	1	-65.7×10^{3}	1
Q1BC	64.5×10^{3}	2	64.5×10^{3}	2
Q2C	-0.524×10^{3}	4	-0.524×10^{3}	4
Q3C	0.495×10^{3}	5	0.495×10^{3}	5
Q4C	-0.010×10^{3}	8	-0.011×10^{3}	8
Q5E	-0.011×10^{3}	7	-0.0105×10^{3}	7
Q5C	0.372×10^{3}	6	0.228×10^{3}	6
β sensitivity S_β				
Transistor (β)	S_β, mV/$\Delta\beta$	Rank	S_β, mV/$\Delta\beta$	Rank
Q1A(200)	−0.141	1	−0.150	1
Q1B(200)	0.141	1	0.150	1
Q2(150)	-0.354×10^{-2}	2	-0.354×10^{-2}	2
Q3(150)	0.262×10^{-2}	3	0.262×10^{-2}	3
Q4(150)	0.169×10^{-2}	4	-0.800×10^{-1}	4
Q5(150)	-0.465×10^{-6}	5	-0.465×10^{-6}	5
Voltage source sensitivity S_V				
Voltage source	S_V, V/V	Rank	S_V, V/V	Rank
V_{D2} (Zener)	0.0178	1	0.00462	1
V_{NEG}	0.0168	2	0.00444	2
V_{POS}	0.0021	3	–	–

To study the second factor affecting the leading edge trigger, the equivalent circuit shown in Fig. 4 was used. This circuit accounts for premature discharging of capacitor C by the nonlinear base current of Q1A and the nonlinear resistance of CR1 as a function of its bias and signal current. In studying the dynamic behavior of the circuit, it was also necessary to simulate mass spectrometer (MS) peaks originating from wide and narrow apertures. For narrow apertures, the MS peaks are nearly triangular; for wide apertures, the waveform becomes trapezoidal. Another required characteristic was the peak-width-to-peak-height relationship. Observed MS peak widths were found to be a maximum of 12 ms for a 10-V peak and half that value for a 100-mV peak. The equations used to describe the two waveforms are:

(1) Trapezoidal waveform.

$$t_1 = \left(\frac{5}{2} + \frac{V_P}{4}\right) \text{ms}$$

$$t_2 = \left[\frac{7}{2} + \left(\frac{7}{20}\right)V_P\right] \text{ms}$$

$$t_3 = \left[6 + \left(\frac{6}{10}\right)V_P\right] \text{ms}$$

(2) Triangular waveform.

$$t_1 = t_2 = \left[3 + \left(\frac{3}{10}\right)V_P\right] \text{ms}$$

$$t_3 = \left[6 + \left(\frac{6}{10}\right)V_P\right] \text{ms}$$

where V_P is the peak value of the input signal expressed in volts.

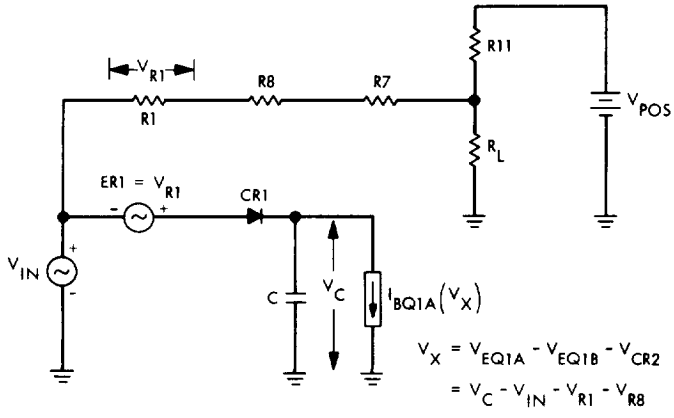


Fig. 4. Leading edge capacitor charging circuit

The purpose of this simulation, which used the *Sceptre* transient analysis program, was to compute the transient trajectory for peak signals ranging from 25 mV to 10 V for both the triangular and trapezoidal waveforms. The trajectory is defined by the parametric equation:

$$V^3_{IN} = f(t)$$

$$V^3_C = g(t)$$

The use of the trajectory information in conjunction with the trigger locus [Table 1; V^1_{IN}, $V^2_C(V_P)$] allows a determination of the actual offset voltage V_{OS} under dynamic conditions. The resulting V_{OS} is equal to or larger than the static offset voltage V^S_{OS} because capacitor C may not have fully charged by the time the input voltage drops sufficiently to cause triggering. The use of the trajectory and trigger locus is illustrated in Fig. 5, and Table 3 summarizes the offsets versus peak input voltage and waveform.

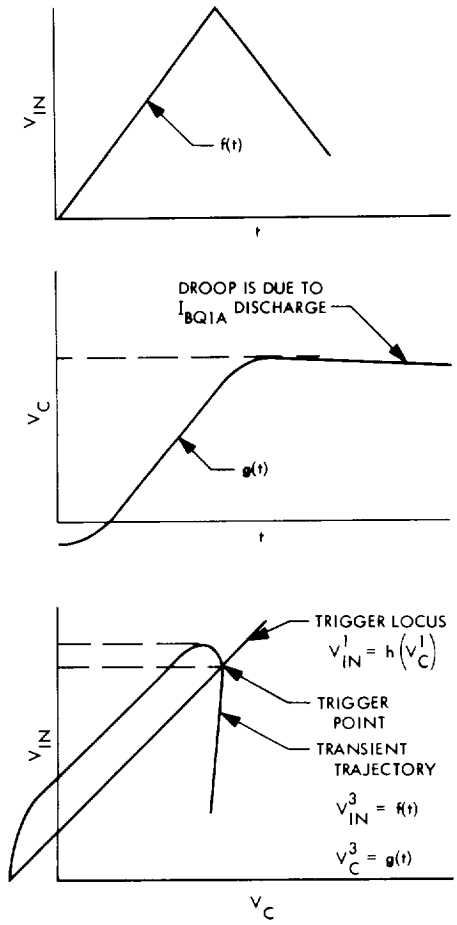


Fig. 5. Example of transient trajectory and trigger locus

Table 3. Dynamic offset voltage as a function of peak input voltage and waveform at 25°C

V_P, V	Triangular			Trapezoidal			
	$t_1 = t_{2f}$, ms	t_{3f}, ms	V_{OS_f}, V	t_{1r}, ms	t_{2r}, ms	t_{3r}, ms	V_{OS_r}, mV
10	6	12	204.2	5	7	12	45.0
5	4.5	9	195.5	3.75	5.25	9	64.3
3	3.9	7.8	184.3	3.25	4.55	7.8	73.0
1	3.3	6.6	147.3	2.75	3.85	6.6	79.1
0.1	3.030	6.06	61.1	2.525	3.535	6.06	51.1
0.05	3.015	6.03	44.8	2.5215	3.5175	6.03	38.8
0.025	3.008	6.015	—	2.5063	3.5088	6.015	—

b. Capacitor discharge. Prior to the leading edge trigger, Q6 is off. When triggering occurs, the PD-A/PWC output V_{OUT} drops to its lower value and the input V_{IN} drops to zero. Q6 turns on and its collector current and the base current of Q1A discharge capacitor C. Thus, the collector current of Q6, I_{CQ6}, largely determines the pulse width t_W.

The current source circuit, shown in Fig. 2, was solved by using PRONE to determine I_{CQ6} and its sensitivity to parameter variation with V_{OUT} in its low state equal to −9.3 V. A summary of results is given in Table 4, where it should be noted that the R_{OUT} referred to is the output resistance of the PD-A/PWC, and that the capacitor voltage V_C was set at 5.0 V.

c. Trailing edge trigger. While capacitor C is discharging, V_{IN} remains at zero and V_{OUT} remains at its lower value. When the capacitor voltage V_C reaches a particular negative value, triggering again occurs and V_{OUT} is restored to its upper value. This action will be called the trailing edge trigger. The computation of the trailing edge trigger is somewhat more simple than the leading edge trigger because only V_C is changing when the threshold is reached.

PRONE was again used to compute the capacitor voltage at which triggering occurs. Because Q5 is just coming out of the saturation region, at the point of triggering, a slightly different feedback setup was used to determine the threshold voltage. Instead of defining the threshold in terms of collector current, it was defined as that capacitor voltage which produces a given collector to ground voltage.

Table 4. Capacitor discharge current source sensitivities (I_{CQ6} = 0.20153 mA at 25°C)

Resistor	Resistor sensitivity S_R	
	S_R, mA/%	Rank
R12	-0.20045×10^{-2}	1
R13	-0.19108×10^{-1}	2
R_{D3}	0.75309×10^{-6}	3
R_{OUT}	-0.14732×10^{-7}	4
Voltage source	**Voltage source sensitivity S_V**	
	S_V, mA/mV	Rank
V_C	0.20303×10^{-10}	3
V_{OUT}	-0.51389×10^{-6}	2
V_{D3}	0.40655×10^{-4}	1
Diode	**Diode junction sensitivity S_D**	
	S_D, mA/mV	Rank
CR6	0.40657×10^{-1}	1
CR7	0.20000×10^{-1}	2
Transistor	**Emitter junction sensitivity S_E**	
	S_E, mA/mV	Rank
Q6	-0.41170×10^{-4}	—
Transistor	**β sensitivity (β = 200)**	
	$S\beta$, mA/$\Delta\beta$	Rank
Q6	0.49873×10^{-5}	—

Since V_{IN} is always zero at the time of the trailing edge trigger, only one solution is required. The solution is:

$$V_C^4 = -0.30303355 \text{ V}$$

Table 5 summarizes capacitor trigger voltage sensitivity to parameter variation.

d. Delay and pulse width. Based upon the results of *Paragraphs a–c*, above, the delay time t_D and the pulse width t_W can now be evaluated and the parameters affecting them discussed. To aid in the discussion, Fig. 6 is provided to define some of the terminology that will be used.

Peak detection delay. It was illustrated in Fig. 5 how the intersection of the transient trajectory and the trigger locus could be used to determine the leading edge trigger. From the illustration, some of the factors which affect the leading edge trigger may also be inferred. For

Table 5. Trailing edge trigger point sensitivities[a]

Resistor	$V_C^{\downarrow}$ resistor sensitivity S_R	
	S_R, V/ %	Rank
R1	-0.477×10^{-4}	6
R2	-0.491×10^{-3}	2
R3	0.158×10^{-4}	8
R4	0.471×10^{-3}	3
R5	-0.171×10^{-4}	7
R6	-0.374×10^{-10}	11
R7	0.250×10^{-3}	4
R8	-0.204×10^{-3}	5
R9	0.470×10^{-1}	1
R10	0.000	12
R11	-0.123×10^{-5}	9
RL	-0.704×10^{-7}	10
Diode	$V_C^{\downarrow}$ diode junction sensitivity S_D	
	S_D, mV/mV	Rank
CR1	0.508×10^{-10}	5
CR2	-0.100×10^{-1}	1
CR3	0.169×10^{-4}	3
CR4	-0.299×10^{-4}	2
CR5	0.214×10^{-9}	4

Transistor (β)	$V_C^{\downarrow}$ β sensitivity S_β	
	S_β, mv/$\Delta\beta$	Rank
Q1A (200)	-0.464×10^{-3}	6
Q1B (200)	0.233	1
Q2 (150)	-0.394×10^{-1}	3
Q3 (150)	0.225×10^{-1}	4
Q4 (150)	−0.186	2
Q5 (150)	0.807×10^{-3}	5
Transistor	$V_C^{\downarrow}$ emitter junction sensitivity S_E	
	S_E, mV/mV	Rank
Q1A	1.000	1
Q1B	−1.000	1
Q2	0.010118	3
Q3	−0.0097293	4
Q4	0.010953	2
Q5	0.000297	5
Voltage source	$V_C^{\downarrow}$ voltage source sensitivity S_V	
	S_V, V/V	Rank
V_{D2} (Zener)	−0.01656	1
V_{NEG}	0.3×10^{-6}	3
V_{POS}	0.186×10^{-4}	2

[a]Nominal 25°C trailing edge capacitor trigger voltage $V_4^C = -0.303034$ V.

instance, if the trigger locus is translated to the left or right, triggering occurs at a new intersection. The factors which cause such translation were discussed in the first of the two considerations controlling the leading edge trigger. A second factor that can affect the point of triggering is the rate at which the input voltage drops. In conjunction with the rate at which the input voltage drops, the base current of Q1A which discharges capacitor C must also be considered. The implication of the base current is best illustrated by taking, for example, the extreme case where the base current discharges the capacitor at the same rate the input voltage drops. Under such circumstances, triggering would never occur. The third and last factor which will be considered is the capacitor charging waveform. Ideally, the waveform of the rising edge of V_{IN} and V_C would be identical. This is not the case because

(1) A nonlinear current-dependent voltage drop occurs across CR1.

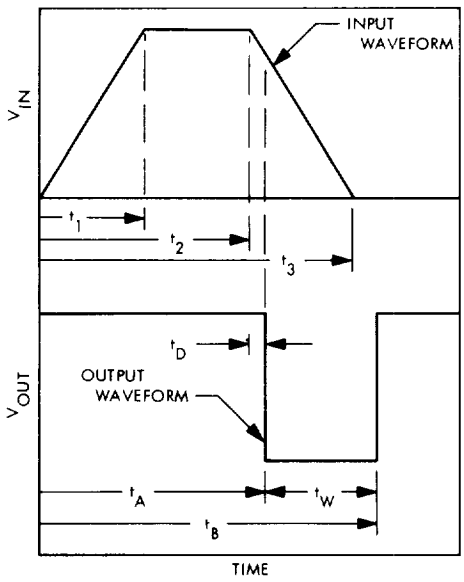
(2) The rate at which capacitor C can charge is limited by the resistance of CR1.

Because diode resistance is inversely proportional to the diode current, small peak input voltages result in larger time constants and correspondingly larger offset voltages. In conjunction with this last point, another related factor should be pointed out. The currents through diodes CR1 and CR2 are

$$I_{CR1} \approx I_{BQ1A} + C\frac{dV_{IN}}{dt}$$

$$I_{CR2} = I_{BQ1B} \approx I_{BQ1A}$$

Therefore, for large peak input voltages with high rates of change, sizable differences can exist between the diode junction voltages of CR1 and CR2 which shift the trigger locus to the right; but because a high rate of change increases the current through CR1, its resistance



t_D PEAK DETECTION DELAY TIME

t_W PULSE WIDTH

t_A TIME OF THE LEADING EDGE TRIGGER

t_B TIME OF THE TRAILING EDGE TRIGGER

t_1, t_2, t_3 DEFINES INPUT WAVEFORM

Fig. 6. Terminology definitions

is substantially reduced, which allows capacitor C to reach steady state more rapidly. This moves the trajectory to the right, thus compensating the waveform-induced delay error.

Figure 7 shows the relationship between delay time and peak input voltage for trapezoidal and triangular input signal voltages. The data used for the curves is given in Table 6.

Pulse width. When the leading edge trigger occurs, the input drops to zero, Q6 turns on, and Q1B turns off. This remains the operating state through most of the discharge period. The capacitor starts at voltage $V_C^3(t_A)$ and discharges at a constant rate until it reaches V_C^5, at which time the trailing edge trigger occurs. The expression for the pulse width t_W is

$$t_W = \frac{C[V_C^3(t_A) - V_C^5(t_B)]}{I_{CQ6} + I_{BQ1A}}$$

The values of pulse widths for trapezoidal and triangular input voltages are plotted in Fig. 7 and tabulated in Table 6.

For all parameters except C, the partial derivative of t_W is

$$\frac{\partial t_W}{\partial P} = \frac{C\left[\frac{\partial V_C^3(t_A)}{\partial P} - \frac{\partial V_C^5(t_B)}{\partial P}\right]}{[I_{CQ6} + I_{BQ1A}]} - \frac{C}{[I_{CQ6} + I_{BQ1A}]^2} \times \left(\frac{\partial I_{CQ6}}{\partial P} + \frac{\partial I_{BQ1A}}{\partial P}\right)[V_C^3(t_A) - V_C^5(t_B)]$$

where P represents any one of the circuit parameters other than C. The terms $\partial V_C^5/\partial P$ and $\partial I_{CQ6}/\partial P$ are summarized in Tables 4 and 5; the partial derivative of $\partial I_{BQ1A}/\partial P$ is small and can be neglected; and most of what was said regarding peak detection delay applies to $\partial V_C^3(t_A)/\partial P$.

e. Input noise sensitivity. For dynamic inputs, the noise threshold starts out as the static offset voltage at zero, which is approximately 17 mv. During charging, the threshold is equal to the static offset plus the additional drop across CR1 due to the charging current. When the input reaches its peak, the capacitor begins to fall, thus decreasing the noise threshold with respect to negative noise spikes. Input noise immunity is further enhanced against low-energy-level noise by the low-pass filtering effect of CR1 and capacitor C.

4. Conclusion

A study of circuit sensitivity to parameter variation shows it to be very stable both with respect to peak detection delay and pulse width. This stability is primarily due to the cancellation effects of the most sensitive parameters. Optimum circuit performance will be achieved if components

Q1A, Q1B

CR1, CR2

Q2, Q3

R2, R4

are matched and the following parameters have small tolerances: R1, R7, R8, R9, R11, R12, D2, D3, C, V_{POS}, and negative voltage V_{NEG}.

The most serious problem of the circuit is the dynamic offset error which is primarily determined by the input signal rate of change. Since for any particular application this error is repeatable, it can be accounted for by calibration.

Table 6. Peak detection delay and analog/pulse width conversion

Triangular					Trapezoidal				
V_P, V	$t_1 = t_2$, ms	t_3, ms	t_D, ms	t_W, ms	t_1, ms	t_2, ms	t_3, ms	t_D, ms	t_W, ms
10	6	12	0.102	2.2814	5	7	12	0.02252	2.3182
5	4.5	9	0.1761	1.1256	3.75	5.25	9	0.04820	1.1560
3	3.9	7.8	0.2396	0.66510	3.25	4.55	7.8	0.07910	0.69090
1	3.3	6.6	0.48610	0.2106	2.75	3.85	6.6	0.21800	0.2263
0.1	3.030	6.06	1.8514	0.02210	2.525	3.535	6.06	1.2930	0.02440
0.05	3.015	6.03	2.6811	0.01430	2.5215	3.5175	6.03	1.9484	0.01570
0.025	3.008	6.015	—	—	2.5063	3.5088	6.015	—	—

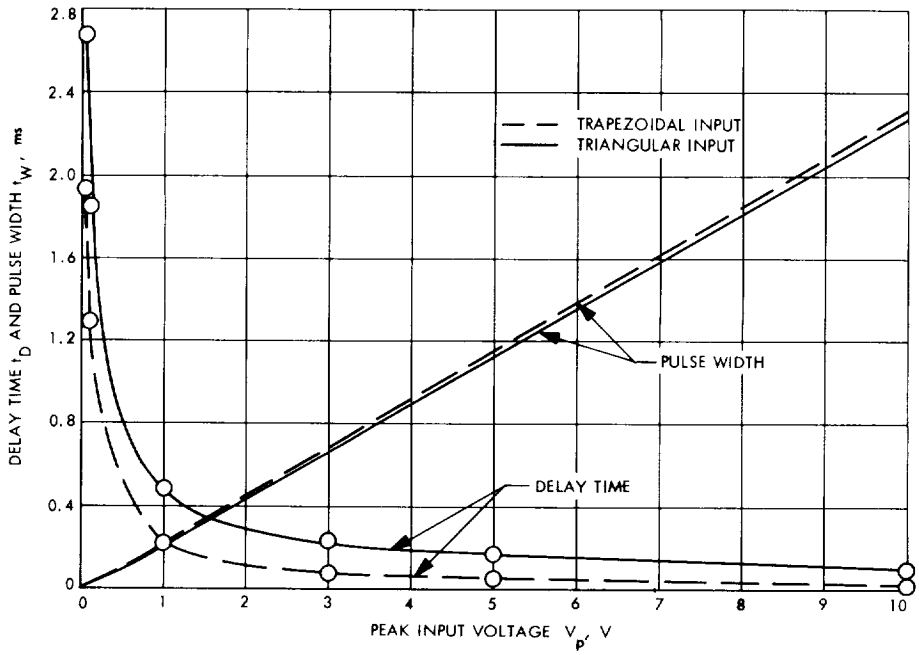


Fig. 7. Delay and pulse width vs peak input voltage

XVI. Bioscience

SPACE SCIENCES DIVISION

A. Soil Studies—Desert Microflora. XV. Analysis of Antarctic Dry Valley Soils by Cultural and Radiorespirometric Methods,[1]

J. S. Hubbard, R. E. Cameron, and A. B. Miller

1. Introduction

In the examination of desert soils, microbiological cultural methods proposed for extraterrestrial life-detection instruments have provided a basis for selecting and testing soil collected in harsh climatic regions of the earth. For this study, Antarctic soils were selected from collections made in McKelvey and Victoria dry valleys in southern Victoria Land during Antarctic austral summer 1966–67 (SPS 37-40, Vol. IV, p. 125, Fig. 6). The samples were collected aseptically from the soil surface and subsurface at seven sites. All samples were stored at temperatures below freezing until used.

2. Soil Properties

All soil physical and chemical analyses were performed on air-dry sieved, or powdered aliquots, by methods previously used for Antarctic soils (SPS 37-44, Vol. IV, pp. 224–236). Some physical and chemical properties of these soils are shown in Table 1. All of the soils are sandy, have pH values above 7.0, and relatively low organic nitrogen and carbon contents, except soil 507. Only soil 537 is very low in soluble salts, and cation exchange capacities are low for all soils except 506, 507, and 508. Moisture content generally increased with depth of sample and with proximity to hard, icy permafrost.

3. Methods

a. Cultural. Microbiological analyses were performed on samples kept frozen until weighed out and inoculated into culture media. One- or 10-g samples were used, and serial dilutions were made for microaerophiles and algae at 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5}. Dilutions for agar plates were at 5×10^{-1}, 10^{-1}, and 5×10^{-3}. The spread plate technique was used to distribute the inoculum on the agar surface. Trypticase soy agar was used to cultivate aerobic bacteria plus actinomycetes, fluid thioglycollate was used for microaerophiles, and cultivation for anaerobes was with trypticase soy agar following the sprinkling of soil on the agar surface. Rose bengal agar with sprinkled soil was used in the attempt to culture fungi, and Thornton's salt medium less organics was used for the algae. All incubations were conducted at 20°C for approximately 4–5 wk, except for the algae. These latter dilution cultures were incubated under Sylvania gro-lux fluorescent tubes at approximately 500 ft-cd light intensity for 3 mo.

[1]Logistic support and facilities for the Antarctic phase of this study were provided R. E. Cameron by the Office of Antarctic Programs, National Science Foundation.

Table 1. Physical and chemical properties of Antarctic soils

Soil	Sample depth, in.	Location	*In situ* moisture content, %	Texture	Electrical conductivity, 10^{-8} mhos/cm² at 25°C	Saturated paste pH	Cation exchange capacity, meq/100 g	Organic nitrogen content, %	Organic carbon content, %
500	Surface 2	McKelvey, near center of valley	1.4	Loamy sand	3360	8.0	3	0.007	0.09
506	Surface 2	McKelvey, near center of valley, 150 ft from site for soil 500	0.84	Sand	2150	7.5	8	0.006	0.09
507	2–6		4.9	Sandy loam	6000	8.2	13	0.057	0.38
508	12		6.9	Sandy loam	3500	8.1	15	0.018	0.19
510	Surface 2	McKelvey, S of Insel Range, SW of Mt. Insel	2.3	Sand	1800	8.0	2	0.008	0.04
511	2–6		1.1	Sand	2150	8.1	2	0.026	0.12
513	Surface 2	McKelvey, N of Olympus Range below Mt. Hercules	2.7	Loamy sand	4000	8.0	2	0.017	0.13
537	Surface 1	Victoria, 150 ft NE of Lake Vida	0.24	Sand	88	8.9	4	0.002	0.02
540	Surface 1	Victoria, ≅ 1 mi	3.5	Loamy sand	3900	8.0	4	0.000	0.05
543	10–12	SW of Lake Vida	4.9	Sandy loam	7000	7.7	11	0.002	0.21
574	Surface 1	Victoria, sand dunes ≅ 1 mi NE of Lake Vida	0.35	Sand	124	9.0	—	0.002	0.02

b. Respirometric. The principle employed in the Gulliver experiment (Ref. 1) was used for detecting biological activity in the Antarctic soils. This procedure involved the measurement of the $^{14}CO_2$ evolved when soils were incubated with ^{14}C-labelled substrates (Table 2). With this method, the metabolic activities can be measured without removing the organisms from their soil environment. The substrates employed were of high specific radioactivity. This permits the detection of the CO_2 produced from the catabolism of a small amount of substrate. Because of the sensitivity of the method, it was necessary to include controls with sterilized soils in order to determine the proportion of the radioactivity which was due to nonbiological decomposition of the substrates. The lowest value considered to be unquestionably positive is twice the background level, i.e., greater than 120 counts/min (net).

Table 2. Metabolic $^{14}CO_2$ production by Antarctic soils[a]

Soil	CO_2 evolved per hour per 300 mg soil, counts/min		
	Untreated	Sterilized	Net[b]
500	673	131	542
506	234	112	122
507	186	181	5
508	94	127	−33
510	683[c]	131[c]	552
511	140	99	41
513	1577	103	1474
537	1008[c]	109[c]	899
540	441	103	301
543	109	140	−31
574	323	121	238

[a]Substrate mixtures contained 144 ng of U-^{14}C-glucose (5×10^4 counts/min) plus 20 ng of U-^{14}C-amino acids (3.9×10^4 counts/min) in 0.3 ml of water.

[b]Counts/min evolved with untreated soil minus counts/min with sterilized soil controls.

[c]Average of three determinations; all other values are the average of determinations.

The measurement of metabolic CO_2 evolution was performed in the apparatus shown in Fig. 1. The soils were handled aseptically and were maintained in a frozen state until immediately before the assay. The reaction was initiated by mixing 0.3 ml of a filter-sterilized solution of the ^{14}C-substrates with 300 mg of soil.[2] The chamber was immediately sealed, and the incubation was carried out

[2]Uniformly labelled ^{14}C-glucose (50 mCi per mmole) was obtained from International Chemical and Nuclear Corp., City of Industry, Calif. Uniformly labelled ^{14}C-amino acids (50 mCi per 33.5 mg) were obtained from New England Nuclear Corp., Boston, Mass. This is a mixture of 15 pure L-amino acids in the same proportions found in algal protein hydrolysate.

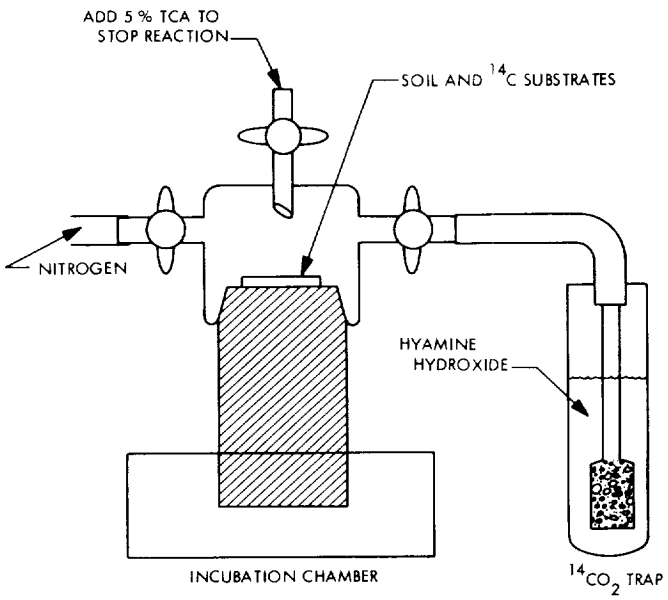


Fig. 1. Apparatus for measuring metabolic $^{14}CO_2$

at room temperature for 1 h. The reaction was stopped by the addition of 0.7 ml of 5% trichloroacetic acid, and the CO_2 which had been evolved was flushed from the chamber with a stream of nitrogen (45 cm^3/min for 30 min). The effluent was flushed through a sintered glass aerator which was submerged in a hyamine hydroxide trap (7 ml of 0.2 M hyamine hydroxide in methanol). The radioactivity in a 1-ml aliquot of the hyamine hydroxide solution was measured in a liquid scintillation counter (Beckman Instruments, model LS-100). The total $^{14}CO_2$ evolution was calculated by subtracting the background radioactivity and correcting for the dilution made from the hyamine hydroxide trap. The control experiments were run with 300-mg aliquots of dry heat sterilized soils using the procedure described for the untreated soils.

4. Results and Discussion

The abundance of microorganisms resulting from culturing on artificial media is shown in Table 3. Compared with most desert soils, the abundances are quite low, showing the absence of anaerobes, fungi, and except for soil 537, also algae. There were no more than 10 bacteria per gram of soil in some of the samples, and only two of the samples, 513 and 537, had abundances approaching 10,000 per gram of soil.

With the exception of soil 506, which was a questionable positive, all the surface soils catalyzed the production of measurable amounts of $^{14}CO_2$ (Table 2). The four soils collected at subsurface levels from three of the sites (507, 508, 511, and 543) did not give a positive metabolic response in this assay. These same four soils also showed the lowest abundance of culturable microorganisms, generally 10 or <10 per gram of soil. Although it is not presently possible to show an exact correlation of the magnitude of $^{14}CO_2$ evolution with the abundance of microorganisms, it should be noted that, in general, the

Table 3. Abundance of microflora in Antarctic soils by cultural methods (per gram of soil)[a]

Soil	Aerobic bacteria + actinomycetes	Microaerophiles (positives at highest dilution)	Anaerobes	Fungi	Algae
500	25	100	0	0	0
506	<10	100	0	0	0
507	10	10	0	0	0
508	<10	10	0	0	0
510	20	1000	0	0	0
511	25	10	0	0	0
513	25[b]	1000	0	0	0
537	8000	10,000	0	0	10
540	55	1000	0	0	0
543	<10	10	0	0	0
574	150	100	0	0	0
Medium	Trypticase soy agar	Fluid thioglycollate	Trypticase soy agar in CO_2	Rose bengal agar	Thornton's medium without organics

[a]All incubations at +20°C for approximately 4–5 wk, except algae for 3 mo.
[b]Actinomycete agar yielded 7,500 pigmented bacteria per gram of soil.

soils with the greatest number of culturable microorganisms showed the greatest $^{14}CO_2$ evolution. For example, as shown in Table 3, soils 513 and 537 contained more culturable microorganisms than the other soils and also showed the highest values for $^{14}CO_2$ evolution.

An anomaly was noted when the data with the Antarctic soils were compared to assays run with soils collected at JPL (unpublished results). Even though the JPL soil contained roughly 1.2×10^7 organisms per gram, the metabolic $^{14}CO_2$ evolved was only 10 times that evolved by soil 513 in which only 7,500 organisms per gram were cultured (Table 3). This would mean that the amount of $^{14}CO_2$ evolved per viable organism in soil 513 is more than 100 times greater than that evolved per JPL-soil organism.

5. Concluding Remarks

Several possibilities are being considered to account for the disparity in the results with the Antarctic and JPL soils. The most obvious explanation is that some of the Antarctic soils contain large populations which do not proliferate on the artificial media used for this study. Alternatively, the organisms of the harsh environment may be adapted to utilize low substrate levels more efficiently than species found in the more favorable environments, i.e., JPL soil. If so, the number of substrate molecules metabolized per Antarctic-soil organism would be higher than that metabolized per JPL-soil organism. Still another possibility is that the Antarctic soils may contain nonviable organisms with active metabolic apparatus to catalyze the CO_2 evolution. Conceivably, the low temperatures of the Antarctic soils could protect these enzyme systems from the thermal denaturation which occurs in warmer soils. Additional cultural and metabolic experiments are in progress to examine these possibilities.

Reference

1. Levin, G. V., et al., *Radioisotopic Biochemical Probe for Extraterrestrial Life*, Third Annual Progress Report, NASA Contract NASr-10, p. III-2. Resources Research, Inc., Washington, D.C., Mar. 30, 1964.

XVII. Fluid Physics

SPACE SCIENCES DIVISION

A. Magnetic Topology and Flux Reconnection,

A. Bratenahl and C. Yeates

Several aspects of the problem of reconnection of lines of force at the X-type, hyperbolic neutral point has been investigated. This work is largely experimental but is supported by theoretical efforts relating it to solar activity. The experience has provided some very valuable insights and revealed some very basic concepts; the neglect or misinterpretation of these concepts has produced some confusion and hindered progress in understanding the problem.

In ordinary hydrodynamics, neglect of viscosity leads to absurd paradoxes unless careful attention is paid to the topology of the flow. For example, exclusion of infinite velocities and velocity gradients introduces such things as boundary layers, wakes, circulation, shocks, and, above all, irreversibility in the form of an essential asymmetry between upstream and downstream (Ref. 1). A very close analogy exists, as might be expected, in hydromagnetics but, unfortunately, this has gained little notice. If resistivity is ignored, a new set of paradoxes is encountered unless it is made certain that the magnetic topology is mathematically free of surface currents. On this basis, the so-called neutral sheet, terminated in branch points, is a wholly inadmissible structure. The significance of the neutral sheets terminated in Y-points has gained considerable interest (e.g., Ref. 2). However, with their elimination, the only critical points available to define a real field are the elliptical 0-points and the hyperbolic X-points. Moreover, if two anti-parallel systems of flux are brought together, the contact line will be broken immediately by an X-point from the very first instant and not delayed (Ref. 3) for some resistive instability to do its work. The topology of a field system is conclusively determined by the number, kind, and geometrical relationship of its zeros, i.e., its critical points. Limitation to 0- and X-points severely limits the variety of possible configurations, and is a very helpful result indeed.

In a region devoid of insulators, the ohmic field at 0-points will destroy flux, while, on the other hand, X-points permit production of new flux through the agency of a suitable velocity field (Ref. 4). It can be shown, in fact, that X-points are a necessary ingredient in any astrophysical dynamo.

The facts discussed above are significant in that: (1) the topology of magnetic and velocity fields is inextricably tied to the requirement of irreversibility (manifested in finite resistivity, finite viscosity) no matter how small these parameters are; (2) the avoidance of paradoxes places topological questions on an equal footing with energetic

questions; (3) the information on the dynamic properties of 0- and X-points must be acquired (with the X-points being of particular importance because of their essential role in the dynamo); and (4) if the detailed component structures, such as spots, flares, and prominences, are to be understood, the whole magnetic system, topologically as well as energetically, must first be understood. Only in this way can the appropriate questions be formulated. The starting point and approach must be basically holistic. Following this procedure, one can expect to derive models of the parts that are consistent with, and responsive to, each other in ways that agree with observations.

This method was used with some success to derive a new dynamically functional quiescent prominence model and a model of Zirin's hourglass flare. The basis of the holistic approach is Babcock's topological synthesis describing a solar magnetic dynamo (Ref. 5), and this work becomes one of its further development and elaboration. At each step of the way, the following questions must be asked: Given the general topology and the kinds of dynamic processes required as the dynamo process evolves, (1) what local structures are implied, (2) where and under what conditions are they to be found, (3) what are they likely to look like, (4) what kind of properties should they exhibit, and (5) can a correspondence be found between these hypothetical structures with those actually seen?

The properties of a flare may not be fully understood, but the experiment reveals precisely what kinds of questions are more important than others. The experiment indicates that the really important question deals with understanding an intricate nonlinear flow process in-the-large. The present day preoccupation of searching for local flows and instabilities that are fast enough is of secondary importance (Ref. 6). The experiment demonstrates that an explosive release of stored magnetic energy does in fact occur at the end point of a self-consistent flow, which leads inevitably toward what is called a critical state. The nature of this limiting state is such that access to it is blocked by the intervention of dissipation and/or resistive instabilities. The question of which instability is thus involved, if any, seems quite academic; for the ultimate instability, if all intervening ones fail, is the temporary breakdown of the plasma condition of space charge neutrality, and with it the failure of high electrical conductivity. The reason is simply that the critical state calls into play the forbidden surface current requiring infinite current density. Carlquist[1] has shown that the maximum current density a plasma can support and still remain a plasma is

$$j_{\max} < ne\left(\frac{kT}{2m}\right)^{1/2}$$

where

n = electronic number density

e = electronic charge

k = Boltzmann's constant

T = temperature

m = electronic mass

The critical state is thus a virtual state, inaccessible in real systems, but of singular importance in controlling the course of events.

The experiment clearly demonstrates a self-consistent conspiracy to set up the forbidden surface current. This general idea is not new (Ref. 4), but what is new is a laboratory demonstration that this critical state could and would be reached in a finite time, were it not for the last-moment intervention of dissipation.

With the experimental device, a collision is produced in the magnetically driven expansion of two cylindrical shock waves in ionized argon. Following collision, the waves merge into an expanding oval, but left behind at the point of collision is an X-type neutral point in a magnetoplasmadynamic flow in a gas of low density and high conductivity. The two parallel driving currents, which increase steadily with time, are ~60,000 A by the time the collision occurs at 3 μs. These currents continue to increase ~200,000 A during the ensuing 7 μs test-time. The condition of rising current requires a very substantial reconnection of flux to take place.

The situation is clarified here by a look at the curl-free potential field about two parallel currents (Fig. 1). The heavy figure eight represents an exceedingly useful concept known as the separatrix (Ref. 7). The separatrix, an integral invariant, can be thought of as a real spacial extension of the hyperbolic neutral point and serves to divide the field system into three subregions on the basis of topological mapping of lines of force about the source currents. Marking a discontinuity in mapping properties (Ref. 8), the separatrix forms the walls of tubular compartments each containing flux of a common topology. The experimental condition of increasing current requires lines of force in regions 1 and 4 to become "severed" so that they can "reconnect" to form new lines in region 2.

[1]Carlquist, P., Report 66-10, Division of Electron and Plasma Physics, Royal Institute of Technology, Stockholm, Sweden, 1967 (to be published in *Phys. Fluids*).

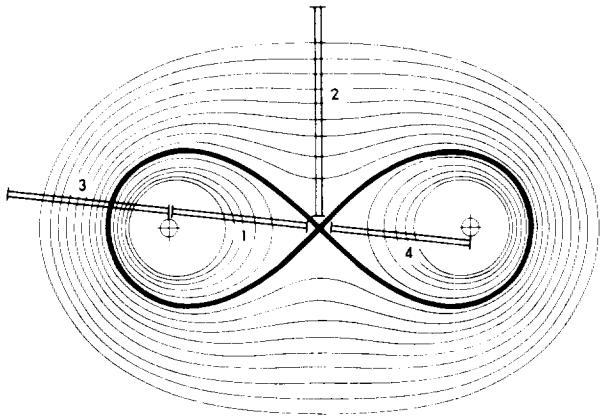


Fig. 1. Curl-free field about two parallel currents

This is a flux transfer process in a topological sense, and studies with magnetic probes and flux probes show that in general this process is unsteady. The numbered double lines in the figure show the position of the flux probes. The flux transfers are, in fact, extremely impulsive (Fig. 2), and the process appears to be describable in terms of a relaxation oscillation. There is reason to suppose that the flux transfers would become even more violent with still higher conductivity simply by delaying significant dissipation until a closer approach is made to the virtual critical state. Thus, in terms of a relaxation oscillation, the time average reconnection rate should depend only weakly on conductivity, if at all, and the time average of the stored magnetic energy should increase as the conductivity is increased.

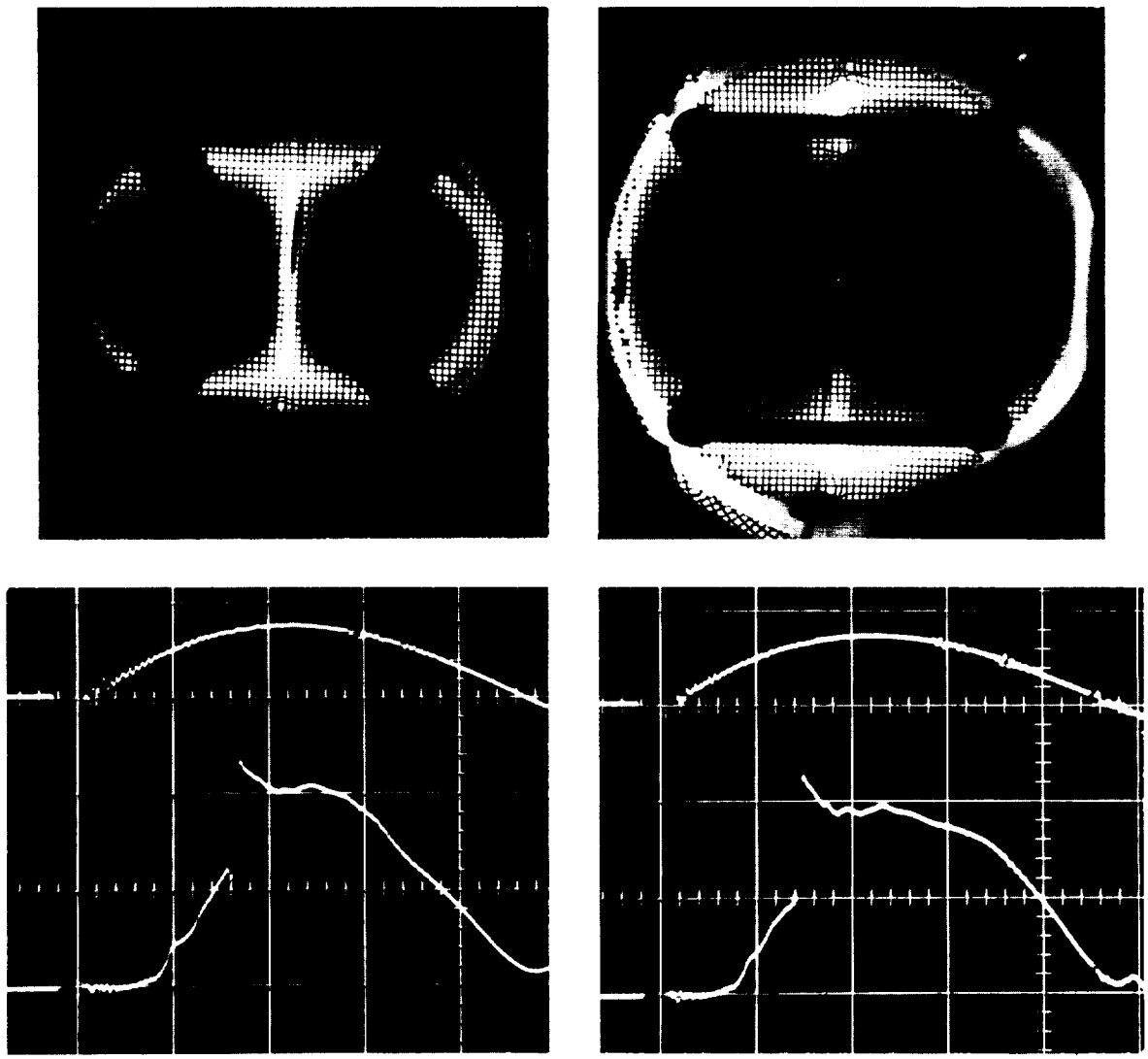


Fig. 2. Oscilloscope traces of large flux transfer event: (upper) main current wave form, (lower) total flux on probe 2 of Fig. 1

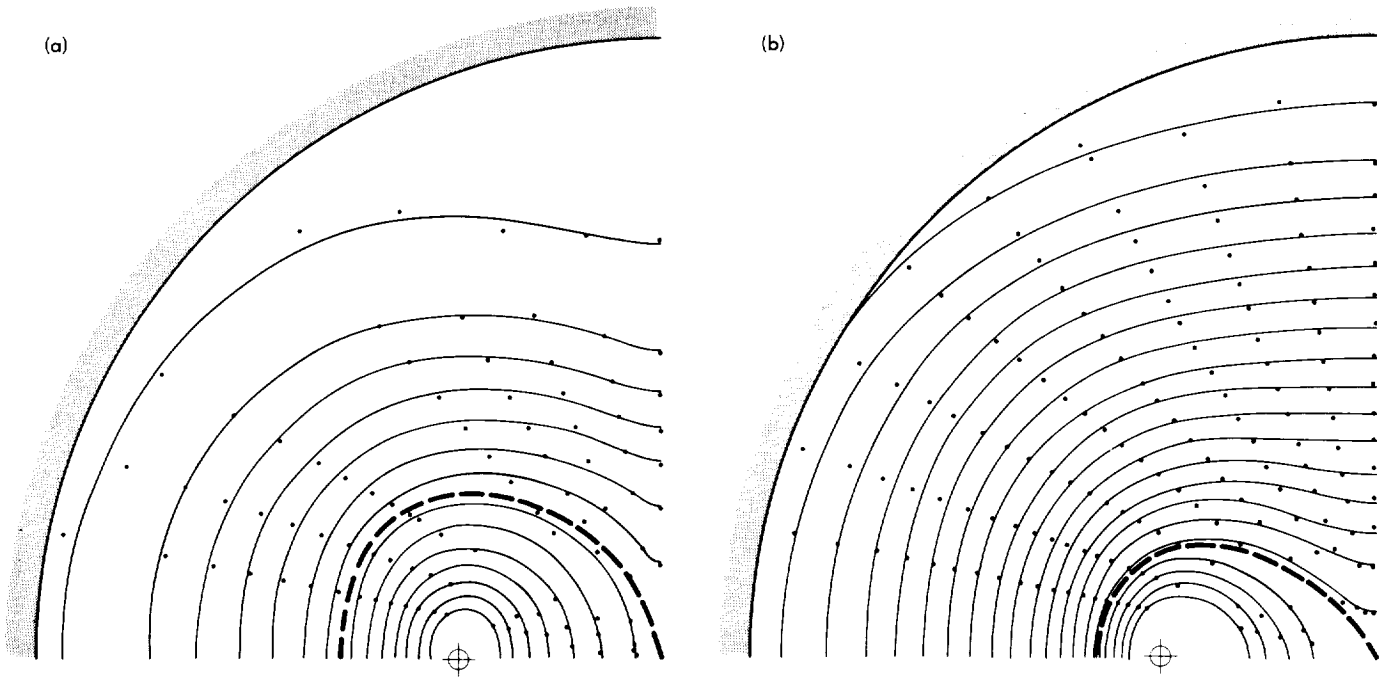


Fig. 3. Magnetic field lines at: (a) 7 μs, (b) 9 μs

Figure 3 shows the status of the lines of force before and after a flux transfer event. A change can be seen in the separatrix (dashed curve), with the distended curve appearing before the event and the contracted curve after. The angle θ this curve makes with the x (horizontal)-axis is related to the current density at the neutral point as

$$j(0,0) = \frac{\mu_0 I}{2\pi a^2}\left(\frac{\tan^2\theta - 1}{\tan\theta}\right), \qquad \theta \to \begin{cases} 0 \\ \\ \dfrac{\pi}{2} \end{cases}, j \to \infty$$

$$\theta \to \frac{\pi}{4}, j \to 0$$

where

μ = permeability of free space

$2a$ = separation of current conductors

I = current

Accompanying the flux transfer, large amplitude magnetoacoustic waves are observed to propagate in all directions from the neutral point. They are expansive in the inner topological regions (x-quadrants) and compressive in the outer (y-quadrants); these waves give a good indication of conditions preceding the event. In the neighborhood of the neutral point, the field develops a curl and the *total current* involved here is completely determined by the amount by which the *distribution of flux* between the inner and outer regions *differs* from the potential distribution. The cross-sectional area through which this current flows, however, is variable, and is determined by the detailed nature of the distortion of the magnetic field that is produced by the flow. The effect is that of a variable resistor whose value is determined by the relative rate of inflow in the x-quadrants to the rate of outflow in the y-quadrants. The direct cause of nonsteadiness in the overall flow is clearly related to the fact that proper coordination of outflow with inflow is impaired by finite signal propagation times. In fact, as a further complication, these times vary in accordance with the previous history of the regions.

In the x-quadrants the Lorentz force is opposed by the pressure gradient; in the y-quadrants it is assisted by the pressure gradient. The result (Fig. 4) is a hyperbolic pinch that is bounded in the x-quadrants by a compressive hydromagnetic shock and in the y-quadrants by a hydromagnetic expansion fan. As neither wave can cross the separatrix, a limit is reached when the finite total current becomes concentrated as a surface current on the separatrix. This is the critical state (step 4). However, the surface current shown in step 5b is forbidden and

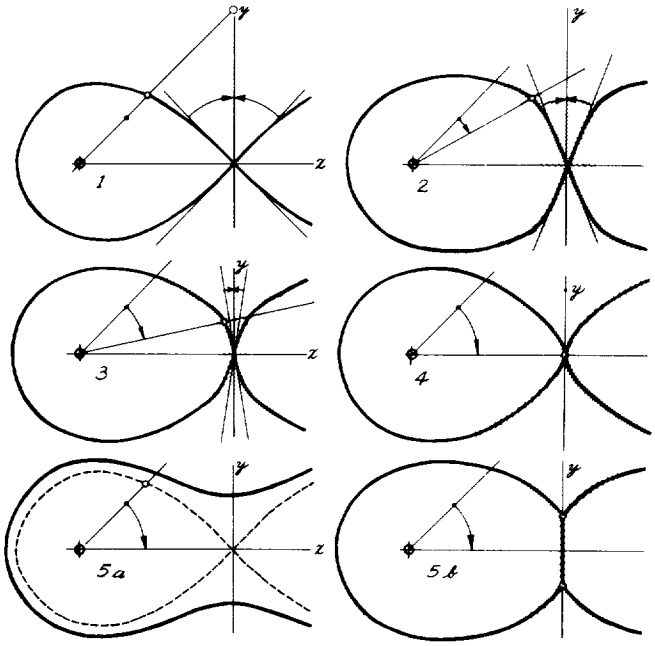


Fig. 4. Proposed behavior of separatrix in relaxation oscillation (hatched area denotes current-carrying region)

actively prevented by any one of a number of possible resistive instabilities that decouple the field from the flowing plasma. Indeed no instabilities need occur at all; the current concentration itself may provide sufficient ohmic electric field in many circumstances. With the decoupling of the field from the fluid, the system quickly relaxes in the direction of the curl-free state (step 5a), and in this transition the stored magnetic energy is released. It is then ready for a repeat performance.

From the manner of the release, it is clear that most of the stored energy is made available for direct plasma and particle acceleration, and very large electric fields have been observed. The readjustment toward the curl-free state spreads at different rates over the entire system. It is therefore clear that: (1) magnetic energy storage is distributed over the entire system; (2) finite propagation times make the problem nonlocal, requiring integral analysis over the whole system; and (3) the control of the current concentration, the variable resistor, is affected by two independent boundary conditions, namely, conditions affecting the inflow and outflow in the x- and y-quadrants, respectively. Petschek's solution (Ref. 9) is incomplete

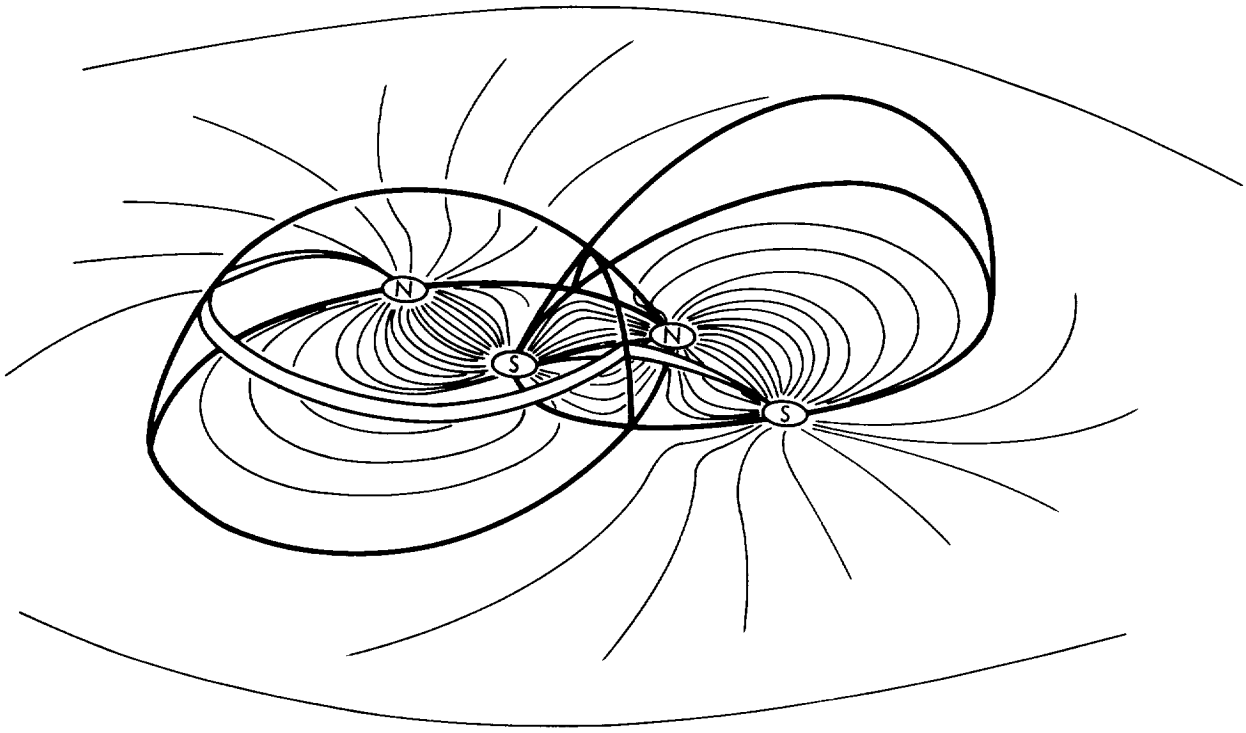


Fig. 5. Separatrix for system of two bipolar sun spots

because its over-simplified geometry precluded adequate treatment of the second boundary condition. It has been demonstrated, for instance, that by suitably impeding the flow in the y-quadrants, the reconnection process becomes steady, and this forms the basis of the new quiescent prominence model.

Such is the observed behavior of the hyperbolic pinch. This work is being continued with an attempt to determine the density and its gradient using Schlieren and hologram methods. The equipment is under development.

If the separatrix has been a help in two-dimensional problems, it is expected to be of much greater help in three-dimensional problems, such as the pair of bipolar sun spot fields shown in Fig. 5. In this case, the separatrix surface decomposes the field into four topologically distinct flux boxes, including the outer space. But in this three-dimensional case something distinctly new comes in. These four chambers have a common line of contact, marked by sharply cornered ridges. This contact line forms an arch joining two genuine X-points lying in the plane of the poles. Contrary to some published opinions (Refs. 7 and 10), this line is not a line of force even though in all probability it satisfies the field line equations. Properly speaking, *this line is a singular integral of these equations corresponding to an envelope solution.* Flux must obviously reconnect somehow all along this line, not just at the X-points at its ends. The true physics of this line must await some future experiment.

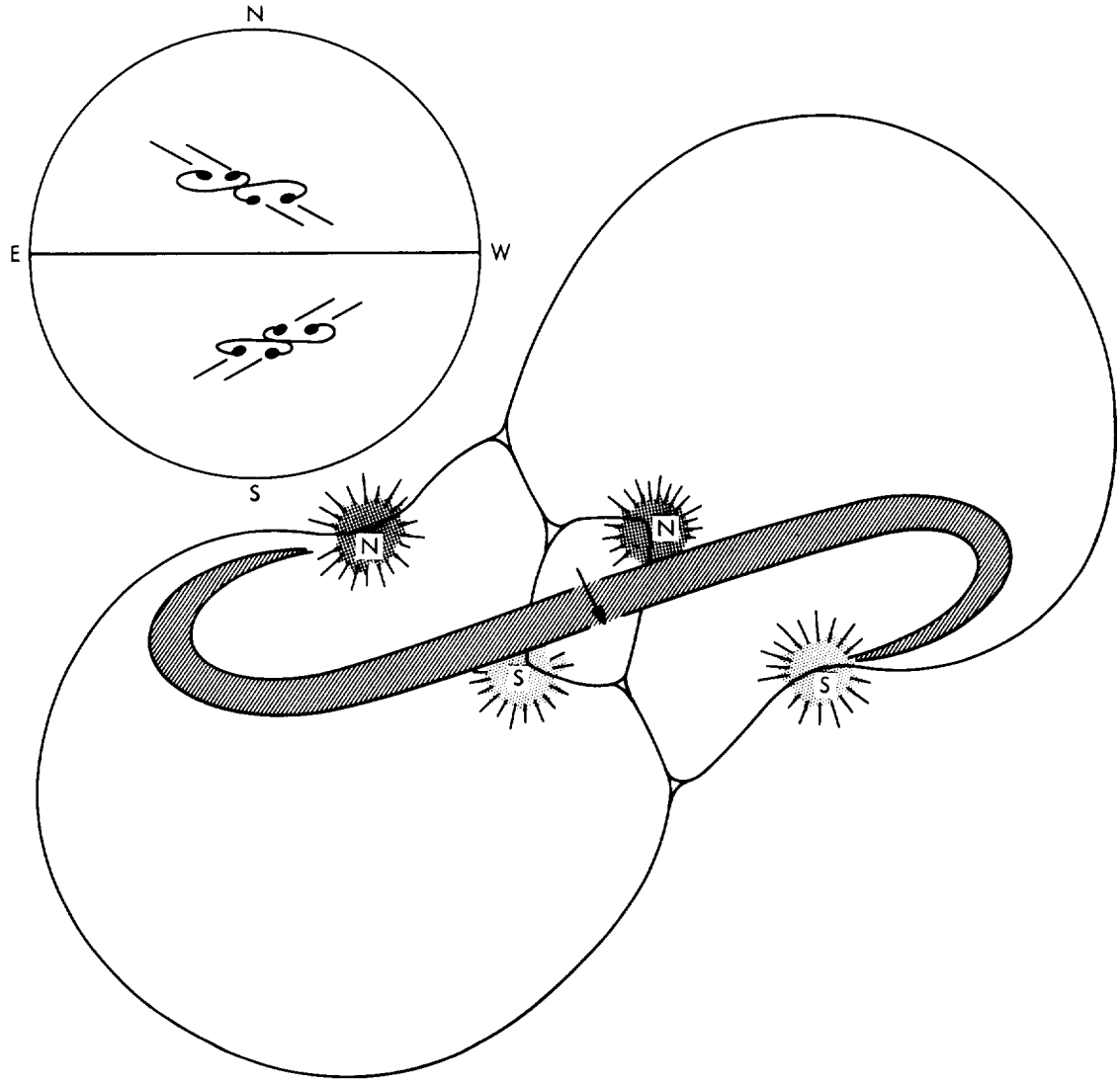


Fig. 6. Orientation rule for S-shape of hourglass flare

Figure 6 shows a possible link between Babcock's theory and Zirin's hourglass flare (Ref. 11), and illustrates the advantage of the holistic approach. If a complex group appears with parallel rows of leader and follower spots, it is clear that new flux coming from below in accordance with Babcock must first enter and overstuff the two egg-shaped boxes. Subsequent transfer of excess flux into the inner and outer boxes releases stored magnetic energy. If, for some reason not understood at present, the flare brightening has the form of the letter S (shown shaded), then it follows that this hourglass flare must show the prescribed parity for the S shown at upper left. Thus far, seven cases have been found supporting this idea and none contradicting it.

References

1. Birkhoff, G., *Hydrodynamics*, Princeton University Press, Princeton, N. J., 1950.
2. Sturrock, P., *Nature*, Vol. 211, p. 695, 1967.
3. Jaggi, R. K., *AAS-NASA Symposium on the Physics of Solar Flares*, NASA SP-50 GPO, p. 419. National Aeronautics and Space Administration, Washington, 1964.
4. Dungey, J. W., *Cosmic Electrodynamics*, pp. 39 and 98. Cambridge University Press, New York, 1958.
5. Babcock, H. W., *Astrophys. J.*, Vol. 133, p. 572, 1961.
6. Gold, T., *Stellar and Solar Magnetic Fields*, p. 390. Edited by R. Lüst. North Holland Publishing Co., Amsterdam, 1965.
7. Morozov, A. I., and Solov'ev, L. S., *Reviews of Plasma Physics*, Vol. II. Edited by M. A. Leontovich. Consultants Bureau, New York, 1966.
8. Sweet, P. A., *Stellar and Solar Magnetic Fields*, p. 377. Edited by R. Lüst. North Holland Publishing Co., Amsterdam, 1965.
9. Petschek, H. E., *AAS-NASA Symposium on the Physics of Solar Flares*, NASA SP-50 GPO, p. 425. National Aeronautics and Space Administration, Washington, 1964.
10. Sweet, P. A., *Nuovo Cimento*, Supplement to Vol. VII, Series X, p. 188, 1958.
11. Zirin, H., *The Solar Atmosphere*, p. 400. Ginn & Blaisdell, New York, 1966.

XVIII. Physics

SPACE SCIENCES DIVISION

A. Enhanced Fluctuations in a Plasma Due to Ion Cyclotron Instability, *C.-S. Wu and J. S. Zmuidzinas*

It is known that the fluctuation fields in a plasma may be resolved into propagating and nonpropagating modes and that under certain circumstances the propagating modes may play a more important role than the nonpropagating modes. For example, Gorbunov and Silin (Ref. 1) have shown that ion oscillations can result in a significant modification of the value of transport coefficients calculated on the basis of the classical binary collision model. Also, Tidman and Dupree (Ref. 2) have demonstrated that collective oscillations can give rise to enhanced bremsstrahlung emission in a nonequilibrium plasma.

One sees from these studies that the essential mechanism generating the said effects is due to enhanced field fluctuations. However, in these and other works, no external magnetic field has been considered. Since in practical cases the plasma is often magnetized, it is desirable to investigate whether enhanced fluctuations are possible in the presence of an applied magnetic field.

The results of a recent study of the spectral density of electric field fluctuations in a magnetized plasma is analyzed. Of particular interest is the contribution of electrostatic oscillations with a frequency close to the ion cyclotron frequency.

The nonequilibrium plasma being considered is described by the following electron and ion distribution functions:

$$F_s(\mathbf{v}) = (2\pi)^{-3/2}\,(u_{sz}u_{s\perp}^2)^{-1}\exp\left[-\frac{(v_z - v_{sD})^2}{2u_{sz}^2} - \frac{v_{s\perp}^2}{2u_{s\perp}^2}\right]$$

where

$$s = e \text{ or } i$$

$$u_{sz}^2 = kT_{sz}/m_s$$

$$u_{s\perp}^2 = kT_{s\perp}/m_s$$

and $v_{eD} \equiv v_D$ is the electron drift speed along the magnetic field $B_0\mathbf{e}_z$. There is no ion drift ($v_{iD} \equiv 0$). The subscripts z and $\perp$ refer to the quantities associated with parallel and perpendicular directions with respect to the magnetic field. It is found that the plasma becomes unstable when v_D exceeds a certain critical value $(v_D)_{crit}$, given by the minimum of the following expression as a function of the wave vector $\mathbf{k}$:

$$x(1 + \Gamma e^{-x^2/2})\,(1 + \delta^{-1})u_{iz}$$

where

$$x = (\omega_r - \Omega_i)/k_z u_{iz}$$

$$\omega_r = (1 + \delta)\Omega_i$$

$$\Omega_i = eB_0/m_i c$$

$$\delta \simeq (T_{ez}/T_{i\perp})e^{-\alpha_i} I_\perp(\alpha_i)$$

$$\Gamma \simeq \left(\frac{m_i}{m_e}\right)^{1/2} \frac{T_{i\perp} T_{ez}^{1/2}}{T_{iz}^{3/2}} \left(1 + \frac{\beta^2 - 1}{1 + \delta}\right) \delta$$

$$\beta = u_{iz}/u_{i\perp}$$

Only the case $v_D < (v_D)_{crit}$ is considered. The resonant contribution to the spectral density of electric field fluctuations is found to be

$$\langle \delta\mathbf{E} \cdot \delta\mathbf{E} \rangle_{\mathbf{k},\omega} \simeq \frac{\delta(\omega - \omega_r)}{\partial \mathrm{Re}\, \epsilon^+/\partial\omega_r} \frac{8\pi^2 u_{ez}^2}{\omega_r - k_z v_D + \omega_0}$$

where

$$\frac{\partial}{\partial\omega_r} \mathrm{Re}\, \epsilon^+(\mathbf{k},\omega_r) \simeq k_e^4/[\beta^2 k^2 k_i^2\, \Omega_i\, e^{-\alpha_i} I_1(\alpha_i)]$$

$$\omega_0 \simeq \left(\frac{m_i}{m_e}\right)^{1/2} \left(\frac{T_{ez}}{T_{iz}}\right)^{3/2} e^{-\alpha_i} I_1(\alpha_i)\, [\omega_r + (\beta^2 - 1)\Omega_i] e^{-x^2/2}$$

In this work, x is always taken to be much greater than unity to ensure weak Landau damping. Consequently, ω_0 is very small. When k_z approaches ω_r/v_D, the second denominator of the spectral density becomes essentially ω_0 so that the spectral density itself is strongly enhanced. Some applications of these results will be discussed in a forthcoming publication.

References

1. Gorbunov, L. M., and Silin, V. P., *Sov. Phys.–Tech. Phys.*, Vol. 9, p. 305, 1964.
2. Tidman, D. A., and Dupree, T. H., *Phys. Fluids*, Vol. 8, p. 1860, 1965.

XIX. Communications Systems Research

TELECOMMUNICATIONS DIVISION

A. Coding and Synchronization Studies: Effect of Quantization on the *Mariner* Mars 1969 Digital Dump Matched Filter, *J. K. Holmes*

1. Introduction

In a previous article (SPS 37-50, Vol. III, pp. 266–272), the digital dump matched filter was analyzed under the assumption that the signals fed into the computer were in analogue form. The actual system, however, must be digital, and hence some degradation is to be expected from this quantization. This article analyzes the effect of a small-gap quantizer inserted at the output of the resistance–capacitance high-pass filter (Fig. 1).

The input sequence, which is composed of the subcarrier (SC) plus (Mod 2) the data (D), together with white noise (N), is multiplied by the subcarrier reference, SC, to produce the noisy data sequence which is fed into the detector. The unit-gain amplifiers are employed to provide isolation stages. The low-pass R_1C_1 (resistance–capacitance) combination, in conjunction with the computer and the analog-to-digital converter, forms the estimate of the symbols every T seconds. The high-pass RC filter is used to "remove" slow drifts that might arise in the amplifiers or in previous circuitry.

2. Effect of the Quantization Noise

In order to analyze the effect of the quantizer on the digital-dump system, it will be assumed that the quantized output is composed of the true output plus quantization noise. To proceed to an expression for the probability of error, the distribution of the quantization noise must be specified. Of course, the distribution in the quantum levels are determined by the distribution of the incoming

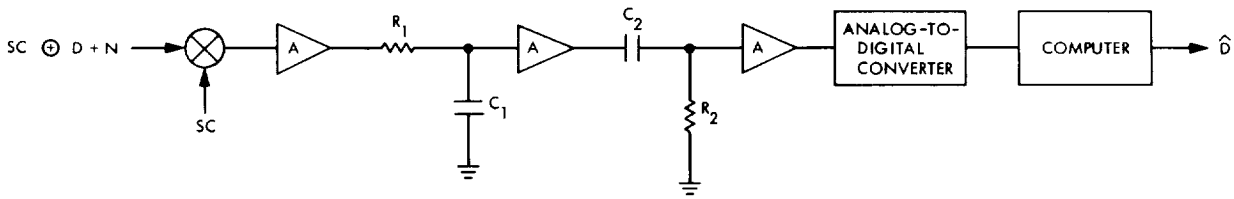


Fig. 1. Digital-dump matched filter with analog-to-digital converter

waveform if we neglect the internal noise of the quantizer. However, this distribution is very difficult to determine, and, in most analyses, the author, in analyzing the problem, will hypothesize a distribution in the gap. In practice, small-gap quantizers (e.g., 14-bit analog-to-digital converters) have been observed[1] to have quantization (conversion) noise, with distributions resembling a quantized-gaussian distribution, with a standard deviation of several quantization levels. This distribution is primarily due to the internal noise of the quantizer. In large-gap quantizers (e.g., 8 level), a uniform distribution is probably a more accurate model. In this article, only a small-gap quantizer is considered, therefore, we hypothesize that the quantizer output can be expressed as

$$X(kT) = Y(kT) + u(kT) \tag{1}$$

where $u(kT)$ is the quantizer noise at time $t = kT$, and, further, that $u(kT)$ is a normally distributed random variable satisfying

$$E[u(kT)] = 0, \qquad E[u(kT)(jT)] = \beta^2 s^2 \delta_{jk} \tag{2}$$

(In effect we are approximating a "discrete gaussian" distribution with a continuous gaussian distribution.) Here β is a constant of the system, s is the quantum-level spacing, and δ_{jk} is the Kronecker delta. It will be further assumed that there is a sufficient number of levels L so that

$$P[Y(kT) \geqq Ls] << P_E$$

where P_E is the average bit-error probability.

The computer forms the decision statistic

$$Z(kT) = X(kT) - X(kT - T)\exp(-T/R_0C_0)$$

which, after using Eq. (1), the above equation becomes

$$Z_k = Y_k - \rho_0 Y_{k-1} + u_k - \rho_0 u_{k-1} \tag{3}$$

We have simplified the notation in Eq. (3), i.e.,

$$Z(kT) = Z_k \quad \text{and} \quad \rho_0 = \exp(-T/R_0C_0)$$

With the assumption that u_k is an independent gaussian random variable, the modification of the previous analysis to include quantization is greatly simplified. Again, fixing the sequence $\mathbf{A}_k$, it is easy to show that $Z_k \mid \mathbf{A}_k$ is a gaussian random variable with its mean and variance given by

$$E[Z_k \mid \mathbf{A}_k] = \frac{\alpha_1}{\alpha_1 - \alpha_2}\mathbf{A}_k(\rho_2 - \rho_1) + \epsilon_k$$

$$V[Z_k \mid \mathbf{A}_k] = \frac{N_0}{4}\frac{\alpha_1^2}{\alpha_1 + \alpha_2}\left[1 + \rho_0^2 - 2\rho_0\left(\frac{\rho_2\alpha_2^2 - (\rho_1 + \rho_2)\alpha_1\alpha_2 + \rho_1\alpha_1^2}{(\alpha_1 - \alpha_2)^2}\right)\right] + \beta^2 s^2(1 + \rho_0^2)$$

$$= \sigma^2 + \sigma_q^2 \tag{4}$$

where ϵ_k represents the second and third terms of Eq. (9) of SPS 37-50, Vol. III, pp. 266–272, σ^2 represents the first term of Eq. (4), and σ_q^2 represents the second term, which exhibits the quantizer variance and, consequently, the effect of the quantizer. It is easy to show, using the same approximation for ϵ_k as before, that the probability of error can be expressed as

$$P_E = \int_{-\infty}^{\infty}\int_{0}^{\infty} \frac{1}{[2\pi(\sigma^2 + \sigma_q^2)]^{1/2}} \exp\left[-\frac{1}{2}\frac{(Z + \gamma + \epsilon)^2}{\sigma^2 + \sigma_q^2}\right] \frac{1}{(2\pi\sigma_\epsilon^2)^{1/2}} \exp\left[-\frac{\epsilon^2}{2\sigma_\epsilon^2}\right] d\epsilon \tag{5}$$

After some algebra, we arrive at the final result

$$P_E = erfc\left[\frac{\gamma}{(\sigma^2 + \sigma_\epsilon^2 + \sigma_\beta^2)^{1/2}}\right] \tag{6}$$

[1]Private communication from J. Layland.

where

$$\sigma_\beta^2 = \beta^2 s^2, \qquad \gamma = \frac{A\alpha_1}{\alpha_1 - \alpha_2}(\rho_2 - \rho_1)$$

and

$$\sigma^2 = \frac{N_0}{4}\frac{\alpha_1^2}{\alpha_1 + \alpha_2}\left\{1 + \rho_0^2 - 2\rho_0 - 2\rho_0\left[\frac{\rho_2\alpha_2^2 - (\rho_1 + \rho_2)\alpha_1\alpha_2 + \rho_1\alpha_1^2}{(\alpha_1 - \alpha_2)^2}\right]\right\}$$

and also

$$\sigma_\epsilon^2 = \frac{A^2\alpha_1^2}{(\alpha_1 - \alpha_2)^2}\left[\frac{1 - \rho_1}{1 + \rho_1}(\rho_1 - \rho_0)^2 + \frac{(1 - \rho_1)(1 - \rho_2)(\rho_1 - \rho_0)(\rho_2 - \rho_0)^2}{1 - \rho_1\rho_2} + \frac{1 - \rho_2}{1 + \rho_2}(\rho_2 - \rho_0)^2\right]$$

Clearly, as $\sigma_\beta^2 \to 0$, the result is identical with the probability of error expression derived assuming no quantization.

B. Combinatorial Communication: Orthogonal Codes and Erasures,[2] *L. R. Welch*

1. Introduction

If a transmitter is using a code with equi-probable, equi-energy code words in a channel with white gaussian noise, the optimal receiver will correlate the received signal with each of the code words and select the code word with the maximum correlation. However, if the maximum is too small, it will be unlikely for the code word selected to be, in fact, the transmitted word. Since sufficient information will be available to compute *a posteriori* probabilities, the receiver could make the alternate decision that an erasure has occurred. This article discusses some relevant statistics.

2. Gaussian Channels with Orthogonal Codes

In a wide-band gaussian channel with white noise, let the transmitted power be P_T, noise power per hertz be N_0, and the bandwidth be B. If D is defined to be the number of degrees of freedom per unit time and A^2 to be the signal energy per degree of freedom, then

$$D = 2B$$
$$A^2 = P_T/2B$$
$$\sigma^2 = N_0B/2B = N_0/2$$

and the channel capacity is

$$C = \frac{1}{2}D\log_2\left(1 + \frac{A^2}{\sigma^2}\right)$$

Since A^2 tends to zero as B tends to infinity, the large bandwidth approximation is

$$C = \frac{1}{2}D\frac{A^2}{\sigma^2\log 2} \tag{1}$$

An orthogonal code is a set, $\{\phi_i \mid i = 1, \cdots, N\}$, of orthonormal functions spanning N degrees of freedom. An encoding is the selection of an integer, i, from 1 to N, and transmission of the function $A(N)^{1/2}\phi_i$. The transmitted energy is then A^2N, or A^2 per degree of freedom, and the information rate is

$$R = D\frac{\log_2 N}{N} \tag{2}$$

The maximum likelihood decoder receives $r(t)$ and selects that i for which

$$X_i = \int r(t)\,\phi_i(t)\,dt$$

is a maximum.

3. Distributions and Probabilities

Let i_0 be the index selected by the transmitter. Then $\{X_i \mid i = 1, \cdots, N\}$ is a set of mutually independent normal random variables. For $i \neq i_0$, X_i has mean zero and variance σ^2 while X_{i_0} has mean $A(N)^{1/2}$ and variance σ^2. The decoder makes an error if

$$X_{i_0} < \max_{i \neq i_0}(X_i)$$

For large N, a good approximate density for $\max(X_i)$ is well known. To simplify its expression, we introduce

[2]Prepared under contract 952108 with the Electrical Engineering Department, University of Southern California.

the new variables

$$Z = \frac{(2\log N)^{1/2}}{\sigma} \max_{i \neq i_0} (X_i) - 2\log N + \frac{1}{2}\log(4\pi \log N)$$

$$Y = \frac{(2\log N)^{1/2}}{\sigma} X_{i_0} - 2\log N + \frac{1}{2}\log(4\pi \log N) \tag{3}$$

Then Z and Y are independent, Y is normal with

$$\mu_Y = \frac{A}{\sigma}(2N\log N)^{1/2} - 2\log N + \frac{1}{2}\log(4\pi \log N)$$

$$\sigma_Y = (2\log N)^{1/2}$$

and, for large N, the density of Z is approximately

$$f_Z(z) = \exp(-z - e^{-z}) \qquad \text{(Ref. 1)}$$

Let

$$M = \max(Y, Z)$$

Then the density of M is

$$f_M(x) = f_Y(x) F_Z(x) + f_Z(x) F_Y(x)$$

where f denotes density and F denotes distribution. A decoding error occurs if $M = Z$. The conditional probability of this event, given $M = x$, is

$$\text{Prob}(M = Z \mid M = x) = \frac{f_Z(x) F_Y(x)}{f_Y(x) F_Z(x) + f_Z(x) F_Y(x)} \tag{4}$$

4. Erasure Decisions

If the receiver has a decision function on observations to accept or reject (erase) the code word of maximum likelihood, then two types of errors can occur. Type I is the acceptance of a wrong code word; Type II is the rejection of the correct code word. A straightforward solution of the variational problem shows that if Type I errors are to be minimized, holding Type II errors fixed, the decision function must reject the maximum likely choice if the *a posteriori* probability that $M = Z$ exceeds a threshold, and accept the choice otherwise. The threshold depends on the Type II probability desired.

In the case where the entire set of observations is used, the *a posteriori* probability is

$$\text{Prob}(M = Z \mid \{X_i\}) = \frac{\exp\left[\frac{\mu_Y}{\sigma_Y^2} \max(X_j)\right]}{\sum_j \exp\left[\frac{\mu_Y}{\sigma_Y^2} X_j\right]}$$

Although this expression may be simple enough for the decoder to use, it is not easy to evaluate Type I and Type II errors for various choices of threshold. A Monte Carlo method seems most feasible but was not attempted.

In the case where

$$\max_j (X_j)$$

is the only observation used, Eq. (4) gives an equivalent function whose behavior can easily be evaluated (see Table 1). It was found that the distributions could be defined in terms of two parameters, $\log_2 N$ and $(C/R)^{1/2}$. For each choice of N, $(C/R)^{1/2}$ was chosen to give a word-error rate of about 1/150 when no erasure rule was used; this rate is the most acceptable for *Mariner*-type video. The threshold was chosen to reduce the error rate (Type I) to about 1/300. Types I and II error rates and erasure rates are shown in Table 1. Notice that when $N = 64$, an erasure rate of triple the initial error rate is needed to reduce the error rate by a factor of 2. Even when $N = 4096$, an erasure rate of 1½ times the initial error rate is needed to reduce the error rate by a factor of 2. It is, therefore, concluded that there is little value in allowing erasures in an orthogonally coded system.

Reference

1. Cramér, H., *Mathematical Methods of Statistics*, p. 375. Princeton University Press, Princeton, N. J., 1946.

Table 1. Types I and II error rates and erasure rates

$\log_2 N$	$(C/R)^{1/2}$	Error rate	Type I rate	Type II rate	Erasure rate
6	1.8	0.0065	0.0032	0.017	0.020
12	1.54	0.0065	0.0032	0.0057	0.009
16	1.46	0.0069	0.0034	0.0043	0.0078
24	1.38	0.0065	0.0032	0.0024	0.0055
30	1.34	0.0066	0.0033	0.0020	0.0052

C. Combinatorial Communication: Sphere-Packing in the Hamming Metric,

R. J. McEliece and H. Rumsey, Jr.

1. Introduction

Let $V_n(2)$ be the n-dimensional vector space over $GF(2)$, with vectors represented as n-tuples of *zeros* and *ones*. The *Hamming metric* $d(x,y)$ is defined to be the number of coordinates in which x and y disagree. If $A = \{a_1, a_2, \cdots, a_M\}$ is a set of M vectors, we define

$$d(A) = \min_{i \neq j} d(a_i, a_j)$$

and

$$\bar{d}(A) = \operatorname{mean}_{i \neq j} d(a_i, a_j)$$

Finally define

$$D(n, M) = \max_{|A| = M} d(A)$$

This article describes a method of obtaining an upper bound on $D(n, M)$ which is always at least as good as the well-known bounds, and which is frequently better. At the same time, the method gives a satisfactory explanation of the relationship between the various known upper bounds on $D(n, M)$. (Hamming and Plotkin, Ref. 1, and Elias[3]). The weakness of the method seems to be that, for the most part, it deals only with the average distance between vectors. Further progress probably awaits a technique capable of dealing more directly with the minimum distance.

2. Derivation

Three theorems are required. In each of these theorems, $A = \{a_1, a_2, \cdots, a_M\}$ is a set of M vectors from $V_n(2)$.

Theorem 1

Let $S_r(x)$ be the sphere of radius r centered at x. Then the mean value of $|S_r(x) \cap A|$ as x varies over $V_n(2)$ is

$$M_r = \frac{M}{2^n} \sum_{k \leq r} \binom{n}{k}$$

Proof

Each a_i appears in exactly

$$\sum_{k \leq r} \binom{n}{k}$$

spheres of radius r, so that

$$\sum_x |S_r(x) \cap A| = M \sum_{k \leq r} \binom{n}{k}$$

Theorem 2 (Plotkin)

Suppose $A \subseteq S_r(x)$, and let the mean distance of vectors in A to x be $\bar{r}$. Then

$$\bar{d}(A) \leq \min\left[2r, 2\frac{M}{M-1}\bar{r}\left(1 - \frac{\bar{r}}{n}\right)\right]$$

Proof

The value $2r$ is obvious. We assume $x = 0$ and arrange the M vectors in the $M \times n$ array (a_{ij}) with column sums s_k. In column k, a pair of entries (a_{ik}, a_{jk}) contribute 1 to $d(a_i, a_j)$ if, and only if, $a_{ik} \neq a_{jk}$. Hence

$$\binom{M}{2}\bar{d} = \Sigma\, d(a_i, a_j) = \Sigma\, s_k(M - s_k) = M\,\Sigma\, s_k - \Sigma\, s_k^2$$

But

$$\Sigma\, s_k = M\bar{r}$$

and by Schwarz' inequality

$$\Sigma\, s_k^2 \geq 1/n\,(\Sigma\, s_k)^2 = M^2 r^2/n$$

so that

$$\binom{M}{2}\bar{d} \leq M^2\bar{r} - M^2\bar{r}^2/n$$

and the theorem follows.

Theorem 3

$$D(n, M) \leq D(n - t, \lceil M/2^t \rceil) \qquad t = 0, 1, 2, \cdots$$

Proof

There must be a set of at least $\lceil M/2^t \rceil$, $\lceil M \rceil$ is the smallest integer $\geq M$, vectors from A which agree on the first t coordinates.

3. Families of Upper Bounds

Using *Theorems 1*, 2, and 3, it is possible to obtain a two-parameter (r and t) family of upper bounds on $D(n, M)$ as follows:

(1) For each r, *Theorem 1* guarantees that we can find a sphere of radius r which contains at least $\lceil M_r \rceil$ vectors from A.

[3]Wyner, A. D., "On Coding and Information Theory," *J. Soc. Ind. Appl. Math.* (to be published).

(2) *Theorem 2* [with $\bar{r}$ replaced by min $(r, n/2]$ then gives an upper bound on the average distance of this subset which is also an upper bound on $d(A)$.

(3) *Theorem 3* allows us to repeat this procedure for the parameters $(n-t, \lceil M/2^t \rceil), t = 1, 2, \cdots$.

The explanation of the relationship between this procedure and the other known bounds is easily stated: If we locate the smallest r for which *Theorems 1* and *2* give any upper bound at all $(M_r > 1)$ and apply the $2r$ part of *Theorem 2*, the result is numerically the same as Hamming's bound. If we apply *Theorems 1* and *2* with the largest allowable r $(r = n)$ to the sequence of pairs $(n-t, \lceil M/2^t \rceil)$, as per *Theorem 3*, the result is Plotkin's bound. (We conjecture that only Plotkin's bound is improved by an application of *Theorem 3*.) Finally, if instead of spheres of radius r we use *shells* of radius $r \leq n/2$, we obtain a somewhat weaker bound. This bound is the same as Elias–Wyner's.

This procedure improves known bounds on $D(n, M)$ for even modest values of the parameters. For example, $D(22, 2^{11}) \leq 6$ is given by the Hamming, Plotkin, and Elias–Wyner bounds, while the procedure of this paper gives $D(22, 2^{11}) \leq 5$. Another interesting example is $D(53, 2^{23}) \leq 18$ (Hamming, Plotkin), ≤ 17 (Elias–Wyner), and $D(53, 2^{23}) \leq 16$ by our methods. It is known (see Footnote 3) that Elias–Wyner's bound is asymptotically better than both the Hamming and the Plotkin bounds. The bound of this paper is, however, not asymptotically better than Elias–Wyner's.

Reference

1. Peterson, W. W., *Error-Correcting Codes*, John Wiley & Sons, Inc., New York, 1961.

D. Combinatorial Communication: A General Formulation of Error Metrics, S. W. Golomb[4]

1. Introduction

In the two decades since Hamming (Ref. 1) described his family of error-correcting codes, a vast and, at times, bewildering literature has developed on the subject of error detection and correction. Central to any formulation of this problem is the notion of an *error metric*. Although there is ample description in the literature of the *Hamming metric* (Ref. 2), occasional reference to the *Lee metric* (Ref. 3), and even some mention of other possible metrics, no attempt has been made to develop a systematic treatment of the general class of metrics that might reasonably correspond to the error patterns encountered in actual communication systems. The purpose of this article is to provide such a treatment.

2. The Alphabet Metric or Anti-Metric

A codeword A may be regarded as vector of dimension k over a q-symbol alphabet: $A = (a_1, a_2, \cdots, a_k)$. A *general error sphere* around this codeword, $S_{hmt}(A)$, may be defined as follows:

A codeword $B = (b_1, b_2, \cdots, b_k)$ is in $S_{hmt}(A)$ if, and only if, the following apply:

(1) B differs from A in no more than h of the k coordinates.

(2) The difference in any one coordinate does not exceed m.

(3) The sum of the differences in all the coordinates does not exceed ℓ.

The parameters m and ℓ require a (one-dimensional) metric on the underlying q-symbol alphabet. For convenience, we will define several possible metrics.

Let the q-symbol alphabet be regarded as consisting of the numbers $0, 1, 2, \cdots, q-1, \bmod q$.

(1) $\rho(a, b) = 1 - \delta_{a,b} = \begin{cases} 0 \text{ if } a = b \\ 1 \text{ if } a \neq b \end{cases}$

(2) $\sigma(a, b) = \min\{(a-b), (b-a), \bmod q\}$

(3) $\tau(a, b) = (b-a), \bmod q$

For both σ and τ, the meaning of $(b-a), \bmod q$, is that integer c, $0 \leq c \leq q-1$ such that $c \equiv (b-a) \pmod q$. In the mathematical sense, $\tau(a, b)$ is not a metric since it fails to satisfy the symmetry law $\tau(a, b) = \tau(b, a)$. (For example, if $q = 10$ and $a = 8$, $b = 7$, we find $\tau(a, b) = 9$ while $\tau(b, a) = 1$. In general, $\tau(a, b) + \tau(b, a) = q$. This may be regarded as an anti-symmetry law, $\tau(a, b) \equiv -\tau(b, a) \pmod q$, and we may call τ an *anti-metric*.) The reason for defining τ is that it corresponds to some of the coding distance ideas considered by several authors, including Shannon (Ref. 4). (In colloquial terms, driving distance between two points may be asymmetric if one-way streets are involved.)

We will assume that any alphabet metric $\mu(a, b)$ which we consider satisfies the remaining metric axioms, namely:

[4]Consultant, Electrical Engineering Dept., University of Southern California.

(1) $\mu(a, b)$ is a non-negative real number for all a and b.

(2) $\mu(a, b) = 0$ if, and only if, $a = b$.

(3) $\mu(a, b) + \mu(b, c) \geqq \mu(a, c)$.

3. The Multi-Dimensional Metric

Suppose $\mu(a, b)$ is any alphabet semi-metric on the q-symbol alphabet Q. That is:

(1) $\mu(a, b)$ is a non-negative real number for all $a, b \in Q$.

(2) $\mu(a, b) = 0$ if, and only if, $a = b$.

(3) $\mu(a, b) + \mu(b, c) \geqq \mu(a, c)$.

Given any two codewords $A = (a_1, a_2, \cdots, a_k)$ and $B = (b_1, b_2, \cdots, b_k)$ of length k over Q, we define the *generalized coding distance function*, $G_{\mu,\alpha}(A, B)$, as

$$G_{\mu,\alpha}(A, B) = \begin{cases} \sum_{i=1}^{k} \mu(a_i, b_i)^{\alpha}, & 0 < \alpha \leqq 1 \\ \left(\sum_{i=1}^{k} \mu(a_i, b_i)^{\alpha} \right)^{1/\alpha}, & \alpha \geqq 1 \end{cases} \tag{1}$$

For any semi-metric μ, we obtain

$$\lim_{\alpha \to 0} G_{\mu,\alpha}(A, B) = H(A, B)$$

equal to the Hamming distance between A and B. Also, if $\mu(a, b) = \rho(a, b)$, we have $G_{\rho,\alpha}(A, B) = H(A, B)$ for all $0 < \alpha \leqq 1$. If $\mu(a, b) = \sigma(a, b)$, we have

$$G_{\sigma,1}(A, B) = \sum_{i=1}^{k} \sigma(a_i, b_i) = L(A, B)$$

equal to the Lee distance between A and B.

That is, the Hamming Sphere of radius r is obtained by setting $h = r$ while leaving m and ℓ unconstrained. (This definition is independent of the alphabet and semi-metric.) The Lee Sphere of radius r corresponds to $\ell = r$ in the σ-metric with no constraints on h and m.

Stein (Ref. 5) has considered certain geometric packing problems that are equivalent to the arrangement of the spheres with $h = 1$ and $\ell = m = r$, with both the σ and τ alphabet (semi-) metrics. Using the σ-metric, with $h = 1$ and $\ell = m = r$, we get the "Stein Sphere of radius r," which is the intersection of the Hamming Sphere of radius 1 with the Lee Sphere of radius r. The Stein Spheres are star-like objects which hug the coordinate axes in k-dimensional space. Using the τ-semi-metric in the same way, we define the "Stein Corner of radius r," also considered by Stein (Ref. 5) as an object with which one may attempt to tile k-dimensional space.

4. A Hierarchy of Error Spheres

The multi-dimensional metric $G_{\mu,\alpha}(A, B)$ is defined by Eq. (1) in the same manner as the Lebesgue metric in the Banach Space L_α. The most interesting Lebesgue metrics are L_0, L_1, L_2, and L_∞. For coding purposes, we have seen that L_0 gives the Hamming metric, irrespective of the alphabet distance function, and L_1 gives the Lee metric when σ is used as the alphabet distance function. A sphere of radius r in the L_∞ metric around the codeword A consists of all codewords which differ from A by as much as r in each coordinate. The measure of difference in a given coordinate is, of course, relative to the particular alphabet metric that has been chosen. It the τ-(semi-) metric is used as the alphabet metric, we will call the corresponding L_∞-spheres "Shannon Spheres," since they generalize a coding problem considered by Shannon (Ref. 4).

If the σ-metric is used as the alphabet metric and the L_∞-metric is used to combine the individual coordinate contributions, we get a hypercube of side $2r + 1$ as our sphere of radius r which we will call the "master sphere of radius r" in k dimensions. The Shannon Sphere of radius r in k dimensions is a hypercube of side $r + 1$.

In Fig. 2, we see a variety of 3-dimensional "spheres" of radius $\leqq 2$ using various coding metrics. These are partially ordered by geometric inclusion, as shown in Fig. 3. (The "Lee Hemisphere" of radius r is defined like the Lee Sphere of radius r, but using τ instead of σ on the alphabet.) All the spheres of Fig. 3, except the Stein Sphere and Stein Corner, are uniquely specified by the alphabet metric and the value of α in Eq. (1).

If we modify the generalized coding distance function $G_{\mu,\alpha}(A, B)$ of Eq. (1) to the non-convex distance function $H_{\mu,\alpha}(A, B)$, defined by

$$H_{\mu,\alpha}(A, B) = \sum_{i=1}^{k} \mu(a_i, b_i)^{\alpha}{}^{1/\alpha}, \qquad 0 < \alpha < \infty \tag{2}$$

then

$$\lim_{\alpha \to 0} H_{\mu,\alpha}(A, B) = S(A, B)$$

the "Stein Distance" between A and B, which is infinite if A and B differ in more than one coordinate. With the distance function H, the Stein Sphere and Stein Corner

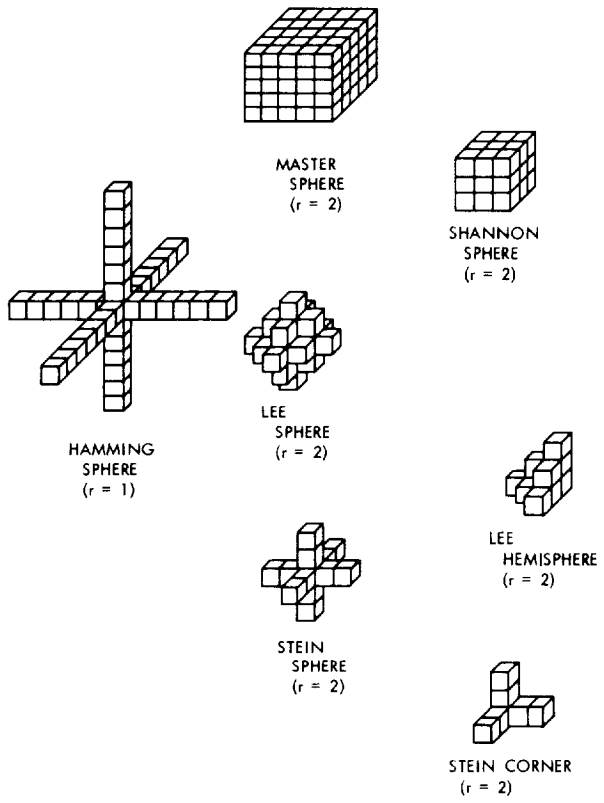


Fig. 2. Three-dimensional spheres of radius ≤ 2 in several different metrics

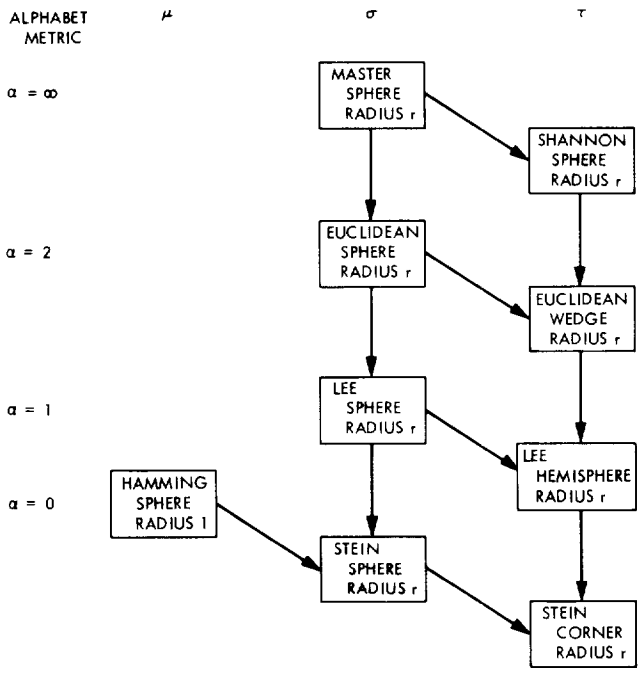


Fig. 3. Hierarchy of "error spheres" (partially ordered by geometric inclusion)

are the spheres of radius r corresponding to $\alpha = 0$ using the alphabet metrics σ and τ, respectively.

5. The Euclidean Case

In classical mathematics, the most important special case of metrics corresponds to $\alpha = 2$, for which value the Banach Space L_α becomes the Hilbert Space L_2, which is self-dual; and normed linear spaces for $\alpha = 2$ satisfy the "parallelogram law" (Ref. 6). Simply stated, $\alpha = 2$ gives us the Euclidean metric, where we have ordinary (Euclidean) spheres of dimension k and radius r around our codeword, in whatever alphabet metric we are using.

For coding purposes, even a Euclidean "sphere" is defined as a set of lattice points. Specifically, the Euclidean sphere of radius r around the point $A = (a_1, a_2, \cdots, a_k)$ consists of all points $X = (x_1, x_2, \cdots, x_k)$ such that

$$\sum_{i=1}^{k} (x_i - a_i)^2 \leq r^2$$

(We have further assumed that σ is the alphabet metric to make this case as familiar as possible.) If $r = 1$, the Euclidean sphere consists of the same lattice points as the Lee Sphere of radius 1. However, for $r > 1$, it is, in general, a difficult problem in number theory just to evaluate the size (i.e., number of lattice points) of the Euclidean sphere. There is, of course, a large literature on this problem involving the representations of integers as sums of squares. The problem of efficient packing of Euclidean spheres is correspondingly difficult.

Coding situations can, indeed, be envisioned for which equal amounts of noise power will produce errors of the same Euclidean magnitude. However, the analysis of any real communications system is likely to lead to a coding metric appropriate to that specific channel, and the Euclidean metric will not be indicated very often. This may be fortunate in view of the evident difficulty of efficient coding for the Euclidean metric.

In Fig. 4, we see how the two-dimensional sphere of radius $r = 3$ increases in size as α goes from 0 to ∞. These

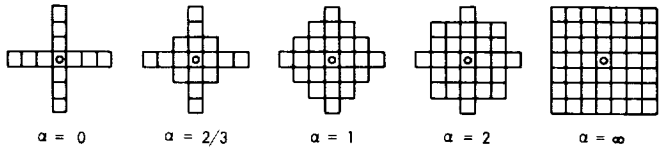


Fig. 4. Two-dimensional spheres of radius 3 for several values of α

"spheres" consist of those points (x, y), as coordinates of the centers of the squares, which satisfy

$$|x|^\alpha + |y|^\alpha \leq 3^\alpha$$

References

1. Hamming, R. W., "Error Detecting and Error Correcting Codes," *Bell Syst. Tech. J.*, Vol. 29, pp. 147–160, 1950.
2. Peterson, W. W., *Error-Correcting Codes*. John Wiley & Sons, Inc.–MIT Press, 1961.
3. Golomb, S. W., and Welch, L. R., "Algebraic Coding for the Lee Metric," in *Conference Proceedings of the Symposium on Error Correcting Codes*, University of Wisconsin, May 1968.
4. Shannon, C. E., "The Zero Error Capacity of a Noisy Channel," *IRE Trans. Inform. Theory*, Vol. IT-2, No. 3, pp. 8–16, Sept. 1956.
5. Stein, S. K., "Factoring by Subsets," *Pac. J. Math.*, Vol. 22, No. 3, pp. 523–541, 1967.
6. Halmos, P. R., *Introduction to Hilbert Space*. Chelsea Publishing Co., New York, 1951.

E. Combinatorial Communication: The Maximum Number of Cycles in the de Bruijn Graph,

H. Fredricksen

1. Introduction

In a recent article (SPS 37-51, Vol. III, pp. 244–250), the problem of disjoint cycles in the de Bruijn graph was introduced. It was conjectured that the maximum number of cycles was attained by the pure-cycling register. The number of cycles defined by the pure-cycling register was given by

$$Z(n) = \frac{1}{n} \sum_{d/n} \phi(d)\, 2^{n/d}$$

where the summation is over all divisors d of the length of the register n, and ϕ is Euler's ϕ-function. This conjecture is discussed in this article, wherein we show that the conjecture is valid in case $1 \leq n \leq 6$. (See SPS 37-49, Vol. III, pp. 295–297 for a discussion of the asymptotic behavior of the conjecture. It was shown that the conjecture is asymptotically correct.) Several algorithms for choosing cycles are investigated and the number of cycles they define is given. We also include some results on the structure of possible counter examples. Proofs of the theorems are generally omitted as the results will be published elsewhere.

2. Z (n) is Maximal for $1 \leq n \leq 6$

If we restrict ourselves to only those truth tables associated with pure cycles, there are $2^{2^{n-1}}$ possible truth tables and a like number of cycle decompositions to consider. For $n = 1, 2, 3, 4, 5$, this number is small enough for a complete search to be made of all the truth tables to determine if the conjecture of the maximality of $Z(n)$ is true or false. For $n = 6$, however, the number of truth tables is too large for a complete search, but we can eliminate this case in another way.

For $n = 1$, there are only two possible truth tables: one where the vector 0 takes on the value 0 (i.e., where $f(x) = f(0) = 0$) and the other where $f(0) = 1$. The first case corresponds to the pure-cycling register. There are two cycles for this register: (0), (1). The other feedback choice yields one cycle, i.e., (01).

For $n = 2$, there are four possible truth tables corresponding to the four independent choices of feedback values that can be associated with the two variables. These cases are:

x_1	x_2	F_0	F_1	F_2	F_3
0	0	0	0	1	1
0	1	0	1	0	1
1	0	1	1	0	0
1	1	1	0	1	0

The pure cycling register is F_0. For F_0 we have $f(0, 0) = 0$ which means the vector 00 maps into the vector 00 corresponding to the cycle (0). The vector 01 maps into 10 since $f(0, 1) = 0$ and 10 maps into 01. This corresponds to the cycle (01). Since 11 maps into 11 we have the cycle (1). We see $Z(2) = 3$. F_1 yields the cycle structure (0) (011). F_2 yields the cycle structure (001) (1). F_3 yields the cycle structure (0011). Thus, the conjecture is valid for $n = 2$.

In SPS 37-51, Vol. III, we gave complete descriptions of the possibilities for cycle decomposition for $n = 3,4,5$. In every case there were no truth tables yielding more than $Z(n)$ cycles.

Since there are approximately 4×10^9 truth tables for $n = 6$, an exhaustive examination of each of them is impossible. We verify the conjecture for $n = 6$ by a separate argument.

First, we write down the cycles generated by the pure-cycling register for $n = 6$, and also the decimal notation for the binary states of the 6-cell register. We then circle certain positions in the cycle decomposition. These positions are the locations of *heavy–light transitions* (see

Table 2. Heavy–light transitions of the pure-cycle decomposition for $n = 6$

Cycles	Cycles in decimal notation
(0)	(0)
(000001)	1, 2, 4, 8, 16, (32)
(000011)	3, 6, 12, 24, 48, (33)
(000101)	5, 10, 20, (40), 17, (34)
(000111)	7, 14, 28, 56, (49), 35
(001)	9, 18, (36)
(001011)	11, 22, (44), 25, (50), 37
(001101)	13, 26, 52, (41), 19, (38)
(001111)	15, 30, 60, 57, (51), 39
(01)	21, (42)
(010111)	23, (46), 29, 58, (53), 43
(011)	27, 54, (45)
(011111)	31, 62, 61, (59), 55, 47
(1)	(63)

Table 2). A number is defined to be heavy if the number has a larger or equal value in the decimal notation than when compared with the vector in the reverse order. If the reverse of a vector is larger than the vector, then that number is said to be light. A heavy–light transition is a situation where either the successor of a heavy number is light or a number is its own successor (see Table 2). We observe from Table 2 that every cycle of the pure-cycle decomposition contains at least one heavy–light transition. We shall show that an arbitrary cycle must contain at least one heavy–light transition.

Theorem 1. Every shift register cycle contains at least one heavy–light transition.

Consider again the set of circled members in the pure-cycle decomposition for $n = 6$ given in Table 2. Since there are 18 circled numbers, there may be as many as 18 cycles. But, if any truth table is to yield more than $Z(6) = 14$ cycles, at least one of the eight circled members that occur paired on a cycle will have to be on a cycle containing no other circled number. We examine each element in turn below to see if they can occur on a cycle which contains no other circled number. An examination of cases shows each circled node is on a cycle with another circled node, hence *Theorem 2.*

Theorem 2. $Z(n)$ is maximum for $n = 1, 2, \cdots, 6$

For $n = 7, Z(n) = 20$ and the number of heavy–light transitions in the pure-cycle decomposition is 34. The methods that disposed of the case $n = 6$ would apply to the case $n = 7$, but the difference between the 34 heavy–light transitions and 20 cycles is too great to allow the necessary computation.

3. Short-Cycle Algorithm

a. Inclusion and exclusion of general cycles. In SPS 37-49, Vol. III, it was assumed that we added the lengths of all cycles discounting the fact that some nodes appeared on more than one cycle. In order for the feedback function to remain single-valued, it is necessary to include only one cycle for each of the 2^n nodes.

We make the reasonable assumption of choosing the shortest cycle available consistent with other choices, which should keep the average cycle length to a minimum. This method suggests we take all cycles that have previously been chosen.

Theorem 3. The cycles of length $\ell \leq [(n + 2)/2]$ contain no overlaps. Table 3 contains the cycle sets resulting from the short-cycle algorithm.

b. Particular results on registers of length 1–7, 8–13, and 17. We note in Table 3 that the short-cycle algorithm results are complete for $n = 1, 2, \cdots, 7$. From the results of *Subsection 2*, we know that no other choice of cycles for $n = 1, 2, \cdots, 6$ will yield more than $Z(n)$ cycles. In *Subsection 4*, we shall investigate other choices of cycles for registers of length $n \geq 7$ in attempting to find a counter example to the conjecture.

In the case $n = 6$, we find the short-cycle algorithm yields 12 cycles while $Z(6) = 14$. We amend the short-cycle algorithm to exclude the three cycles of length 4 which can be chosen. This allows us to choose four cycles of length 5, (00001), (00011), (00111), and (01111), which had been excluded by the cycles of length 4, and also to choose four cycles of length 6, (001011), (001101), (000101) and (010111). With one cycle of length 10, we have a total of 14 cycles for $n = 6$. This increase of two cycles gives hope for finding a counter example in *Subsection 4.*

For $n = 6$, 8–13, 17, the short-cycle algorithm yields less than $Z(n)$ cycles. If the computer had been allowed to find all the cycles in each case, the time required would have been excessive. The computer was instructed on a second pass to stop generating cycles after it was clear that the algorithm would fail to produce $Z(n)$ cycles. For these n, Table 4 gives the bound on the number of cycles produced by the algorithm and the length of the cycle at which time it was clear $Z(n)$ cycles would not be achieved. The cycle length at halt is in no sense monotone, and its asymptotic behavior is not clear. Because it can be said that the algorithm will fail to produce $Z(n)$ cycles at such small values compared to n, it appears that it will not ultimately produce a counter example.

Table 3. Cycles from the short-cycle algorithm

n \ ℓ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	20	22
1	2																			
2	2	1																		
3	2	1	—	1																
4	2	1	2	—	—	1														
5	2	1	2	1	—	1						1								1
6	2	1	2	3	—	1	2													
7	2	1	2	3	4	2	—	3							2				1	
8	2	1	2	3	6	5	—	1	—	6	1	—	—							
9	2	1	2	3	6	7	6	2	4	7	4	4	—	2						
10	2	1	2	3	6	9	12	9	—	11	—	6	6	4	3					
11	2	1	2	3	6	9	16	18	14	11	4	16	6	8	7	12				
12	2	1	2	3	6	9	18	26	34	21	10	24	16	19	10	9	6			
13	2	1	2	3	6	9	18	28	46	56	36	20	36							
17	2	1	2	3	6	9	18	30	56	97	174	287	532	479	226	226	548	286		

Table 4. Bounds on the number of cycles produced by the short-cycle algorithm

n	Z (n)	Bound	Cycle length at halt
6	14	12	7
8	36	32	10
9	60	≤58	13
10	108	≤98	12
11	188	≤177	14
12	352	≤320	14
17	7712	≤7616	18

4. Other Algorithms

a. Choose cycles of length $n-1$. The results of *Subsection* 3 indicate that the short-cycle algorithm will not provide a counter example. The pure-cycling register yields all the cycles of length k for $k|n$. Another approach, intermediate to the short-cycle and the pure-cycle algorithm, is the following:

(1) Take all cycles of length $n-1$.

(2) Take all smaller cycles not excluded.

(3) Take smallest remaining cycles.

Theorem 4. This algorithm yields $\sim \frac{3}{4} Z(n)$ cycles as $n \to \infty$. The proof of this theorem is accomplished by showing several intermediate results.

Lemma 1. No $(n-1)$ cycle can exclude another.

Lemma 2. Any t-cycle, where $t|(n-1)$, is disjoint from any v-cycle if $v|(n-1)$.

An examination of the truth table imposed by this choice of cycles yields the result.

For short registers, the algorithm does somewhat better than its asymptotic behavior. For $n = 1, 2, 3, 4, 5$, and 7, the algorithm produces $Z(n)$ cycles. For $n = 1, 2, 3, 4$, the same cycle sets are produced as were produced by the short-cycle algorithm. For all other n, this algorithm produces different cycles than did the short-cycle algorithm. For $n = 5$, we have the three cycles of length 4, two cycles of length 1, and one cycle of length 2 from the pure cycling half of the truth table. The other nodes are on two cycles of length 8 for the total of eight cycles.

The $n = 7$ case produces the 14 cycles for the pure-cycle half of the register corresponding to $Z(6) = 14$. The complementing half produces one cycle of length 4 and five cycles of length 12 for a total of 20 cycles.

We can extend the above results by starting with all cycles of length $n-2$, $n-3$, etc., and continuing with the rest of the algorithm. We produce $Z(n-2)$ cycles in the case of length $(n-2)$ cycles. This uses 2^{n-2} nodes of the diagram. We can artificially impose a complement restriction on another 2^{n-2} of the nodes in the same way as was naturally imposed in the length $(n-1)$ algorithm producing cycles of length $2(n-2)$. At this point, the remaining 2^{n-1} nodes are on cycles of length $4(n-2)$. This method produces approximately

$$\frac{2^{n-2}}{n-2} + \frac{2^{n-2}}{2(n-2)} + \frac{2^{n-1}}{4(n-2)} = \frac{2^n + 2^{n-1} + 2^{n-1}}{4(n-2)} =$$

$$\frac{2^n}{2(n-2)} \sim \frac{1}{2} Z(n) \text{ cycles}$$

If we start with cycles of length $(n-3)$, we produce approximately

$$\frac{2^{n-3}}{n-3}+\frac{2^{n-3}}{2(n-3)}+\frac{2^{n-2}}{4(n-3)}+\frac{2^{n-1}}{8(n-3)}=$$

$$\frac{2^n+2^{n-1}+2^{n-1}+2^{n-1}}{8(n-3)}=\frac{5}{16}\frac{2^n}{(n-3)}\sim\frac{5}{16}Z(n)\text{ cycles}$$

For cycles of length $(n-4)$, we produce approximately

$$\frac{2^{n-4}}{n-4}+\frac{2^{n-4}}{2(n-4)}+\frac{2^{n-3}}{4(n-4)}+\frac{2^{n-2}}{8(n-4)}+\frac{2^{n-1}}{16(n-4)}=$$

$$\frac{2^n+2^{n-1}+2^{n-1}+2^{n-1}+2^{n-1}}{16(n-4)}=$$

$$\frac{3}{16}\frac{2^n}{(n-4)}\sim\frac{3}{16}Z(n)\text{ cycles}$$

In general, this method yields approximately

$$\frac{k+2}{2^{k+1}}Z(n)$$

cycles if we start with cycles of length $(n-k)$.

b. Choose cycles of length $n-2$, $n-3$, etc. If the artificial complement restriction is not imposed, other cycle distributions will be produced if we start with the cycles of length $n-2$, $n-3$, etc. For $n\leq 6$, the short-cycle algorithm produces the same cycles since the length of the longest unrestricted cycles in the short-cycle algorithm was $[(n+2)/2]$.

In SPS 37-49, Vol. III, and *Subsections 3* and *4*, we gave algorithms that yielded certain cycle structures. The methods used in these algorithms required that the shortest cycle not explicitly ruled out by having a non-void intersection with the set of cycles previously chosen be included in the set. For cycles of length not exceeding the register length, the set of cycles chosen is unambiguous. For cycle lengths exceeding the register length, there exist cycles not excluded by any shorter cycles, but having non-void intersection with other cycles of that length. The choice of cycles to be included is arbitrary for these lengths. A hybrid method can be used to consider all choices of cycles for lengths exceeding the register length.

5. Structure of Possible Counter Examples

In this section we shall give some results on the structure of possible counter examples.

a. Weight of counter example ≥ 7. We handle the proof with 3 lemmas.

A counter example of weight 7 could be formed from four cycles as follows:

(1) Join cycle A to B, B to C, and C to D (the weight is now equal to 3).

(2) Cycles A and B can be split once, as can B and C and C and D.

(3) This yields 4 cycles and a truth table of weight 6.

(4) Cycle A was a cycle containing k ones.

(5) Then cycle B contained $k\pm 1$ ones.

(6) Cycle C has $k\pm 2$, or k, ones; then cycle D could have $k\pm 1$ ones and could contain, with A, a vector-successor pair.

b. Minimum-join–maximum-split path. Suppose we do have a counter example to the conjecture for some n. We find one with a truth table of minimum weight, $(2\ell+1)$. This counter example must have $Z(n)+1$ cycles since, if it had more cycles, we could decrease the weight of the truth table and have a counter example of smaller weight. There are $(2\ell+1)!$ paths from the position $[2\ell+1, Z(n)+1]$ to the position $[0, Z(n)]$. As we decrease the weight of the truth table, the number of cycles changes. Choose one of the paths that stays closest to the $Z(n)$ line, i.e., consider the first set of joins of cycles (as we begin to decrease the weight of the truth table) and choose a path having the minimum number of joins before one of its cycles is capable of splitting. If more than one path has the minimum number of joins, choose one that has a maximum number of splits after the initial set of joins. Continue the process until all $2\ell+1$ ones in the truth table have been changed to zeros. We seek to test the conjecture that the truth tables lying on the minimum-join–maximum-split path formed by the above procedure never yield less than $Z(n)-1$ cycles and give a set of lemmas on the general nature of the minimum-join–maximum-split path. For the purposes of the lemmas, we shall consider the path as the weight of the truth table increases. Then the joins will be splits, and the splits will be joins.

Lemma 3. As the weight of the truth table increases, the set of all splits occurring at any point of the path comes from exactly one cycle.

Lemma 4. As the weight increases, the set of $t>1$ splits can be preceded by, at most, one join.

Lemma 5. The minimum-join–maximum-split path cannot miss exactly one position anywhere on the $Z(n)$ line (see Fig. 5a).

Lemma 6. The minimum-join–maximum-split path cannot miss exactly two positions at the end of the path (see Fig. 5b).

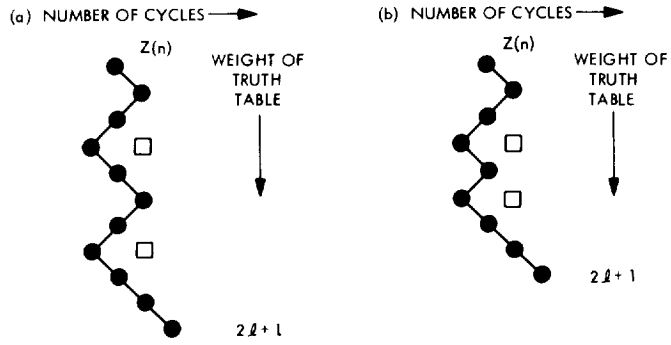


Fig. 5. Z (n) line showing (a) miss 1 position on the minimum-join–maximum-split path, and (b) miss 2 positions on the minimum-join–maximum-split path

F. Information Processing: Prediction With Piecewise Linear Correlation Functions, *I. F. Blake*

1. Introduction

This article considers one aspect of the problem of predicting a continuous-time parameter random process. It is shown that for a particular class of piecewise linear (PL) correlation functions, the optimal (minimum mean-square error) linear predictor utilizes only a finite number of samples from any finite observation interval.

2. Prediction With PL Correlation Functions

Consider a continuous PL correlation function that is zero outside of a finite interval (which, without loss of generality, is assumed to be $[-1, 1]$). Such a function may be represented by letting $\phi(\tau) = \max(0, 1 - |\tau|)$ and

$$\rho(\tau) = \sum_{i=1}^{n} p_i \phi(\tau/a_i) \tag{1}$$

where

$$\sum_{i=1}^{n} p_i = 1, \qquad a_i < 1, \qquad a_n = 1$$

Clearly, $\mathbf{a} = \{0, a_1, \cdots, a_{n-1}, 1\}$ is the set of points at which $\rho(\tau)$ changes slope. Denote the interval $[0, T]$ over which the process has been observed by π, and suppose an estimate of the process at $T + \alpha$ is desired. The point set Ω, which will be called the sampling set, is given by the following:

Definition. The set Ω is such that

(1) $\mathbf{a} \cap \pi \in \Omega$, $(T - \mathbf{a}) \cap \pi \in \Omega$, and $(T + \alpha - \mathbf{a}) \cap \pi \in \Omega$.

(2) If $t_i \in \Omega$, then $\{(t_i + \mathbf{a}) \cup (t_i - \mathbf{a})\} \cap \pi \in \Omega$.

Note that Ω is a function of $\mathbf{a}$, T, and α. It is apparent, in many cases, that the cardinality of the set Ω is finite ($[\Omega] < \infty$). Conditions on $\mathbf{a}$ and π for this to occur are made precise in the following theorem.

Theorem 1. Assume π to be finite. Then a sufficient condition for $[\Omega] < \infty$ is that all break points be rationally related. A necessary condition for this is that if any two break points a_i and a_j are in π, and their sum $a_i + a_j$ is also in π, then they must be rationally related.

Proof. First assume that $a_i = i/n$, n a positive integer. Then, for arbitrary T and α, the set Ω will consist of the points

$$\{0, 1/n, 2/n, \cdots, T, T - 1/n, T - 2/n, \cdots,$$

$$(T + \alpha) - (1/n), (T + \alpha) - (2/n), \cdots\} \cap \pi$$

and, clearly, this set will contain finitely many points. Omitting a break point, say a_j, from $\mathbf{a}$ will generate a set Ω' with no more points than the above Ω, i.e., $\Omega' \subseteq \Omega$. It follows that if $a_i = m_i/n_i$ (m_i, n_i are integers), then $[\Omega] < \infty$ since $n_1 n_2 \cdots n_n$ forms a greatest common denominator.

For the converse, suppose that the two break points are not rationally related and $a_i, a_j, a_i + a_j \in \pi$ and $a_i < a_j$. Consider the number of points that must be in Ω from the interval $[0, a_i + a_j]$. Begin the development by adding a_i to a_i. If the sum is in the interval $[0, a_j]$, add a_i again. If the sum is in $(a_j, a_i + a_j)$, subtract a_j. Continuing in this way, it is clear that there will be an infinite number of points in $[0, a_i + a_j]$ (and hence in Ω) of the form $ma_i - na_j$ where m and n are integers. It remains to show that no two of these points are identical. Suppose two of the points were identical, i.e., there exist integers m, n, m', n' such that $ma_i - na_j \in \Omega$, $m'a_i - n'a_j \in \Omega$ and that

$$ma_i - na_j = m'a_i - n'a_j \tag{2}$$

or

$$a_i(m - m') = a_j(n - n') \tag{3}$$

This, however, implies that a_i and a_j are rationally related which is contrary to the assumption. Therefore, no two of the points generated in this manner are identical, and, under the assumptions stated, Ω will not have finite cardinality.[5]

The principal result of this article is now given in the following theorem.

Theorem 2. For any real random process $x(t)\ t \in \pi$ (π finite), with a continuous PL correlation function $\rho(\tau)$ having a break vector **a** that generates the set Ω with finite cardinality, the linear least mean-square error predictor depends only on the points

$$t_i \in \Omega, \qquad i = 1, \cdots, M$$

Proof. Assume that the M points t_i in Ω are ordered, i.e., $0 = t_1 < t_2 < \cdots < t_M = T$. It is only necessary to show that

$$E\left\{\left[x(T+\alpha) - \sum_{i=1}^{M} d_i x(t_i)\right] x(t)\right\} = 0 \qquad t_i \in \Omega \tag{4}$$

and all $t \in \pi$ where the constants d_i are chosen such that

$$E\left\{\left[x(T+\alpha) - \sum_{i=1}^{M} d_i x(t_i)\right] x(t_j)\right\} = 0 \qquad j = 1, \cdots, M \tag{5}$$

Denote the closed interval

$$[t_K, t_{K+1}] \text{ by } J_K, \qquad K = 1, \cdots, M-1$$

and consider the quantity y_t where

$$y_t = E\left\{\left[x(T+\alpha) - \sum_{i=1}^{M} d_i x(t_i)\right] x(t)\right\} \qquad t \in J_K \tag{6}$$

for some arbitrary but fixed K. It is claimed that for all $t \in J_K$, the right-hand side of Eq. (6) is given by an equation that is *linear* in t, and that is the *same* equation for the whole interval J_K. To see this, it need only be shown that the expression for $\rho(|t - t_i|)$ does not "pass through" any of its break points as t varies over J_K for each $t_i \in \Omega$. This is clear because condition (2) in the definition of Ω states that if $t_i \in \Omega$, then $\{(t_i + \mathbf{a}) \cup (t_i - \mathbf{a})\} \cap \pi \in \Omega$ and, hence, if a break point of $\rho(|t - t_i|)$ occurred in J_K, that point would also occur in Ω. Therefore, t_K and t_{K+1} would not have been adjacent, as assumed, and for each $t_i, i = 1, \cdots, M$, and for all $t \in J_K$

$$-d_i \rho(|t - t_i|) = \beta_{iK} + \gamma_{iK} t \tag{7}$$

Thus, Eq. (6) may be written as

$$y_t = \beta_K + \gamma_K t \tag{8}$$

where

$$\beta_K = \sum_{i=0}^{M} \beta_{iK} \qquad \gamma_K = \sum_{i=0}^{M} \gamma_{iK}$$

and

$$\beta_{0K} + \gamma_{0K} t = \rho(T + \alpha - t)$$

But from Eq. (5), it is seen that

$$y_{t_K} = \beta_K + \gamma_K t_K = 0 \tag{9}$$

and

$$y_{t_{K+1}} = \beta_K + \gamma_K t_{K+1} = 0 \tag{10}$$

Thus, $\beta_K = \gamma_K = 0$ and Eq. (6) is zero for all $t \in J_K$. Since K was chosen arbitrarily, the same is true for all the intervals and, hence, for π.

3. Comments

A simple consequence of *Theorem 2* is that if it is desired to predict a given process α into the future where $1/\alpha = n =$ integer, the break points of the PL correlation function are evenly spaced α apart, and T is some multiple of α, then the best linear predictor uses only uniformly spaced samples on π.

For the particular case where

$$\rho(\tau) = \phi(\tau) = \max(0, 1 - |\tau|)$$

the prediction formula may be stated in general terms.

[5]The author wishes to thank Dr. L. L. Welch for supplying the last part of this proof.

For $\alpha < 1$, the formula is

$$\hat{E}[x(T+\alpha) \mid x(t), t \in \pi, N-1 \leq T < N] = \left[b - \frac{1}{N+1}\right]\left[x(0) + \frac{N-1}{N}x(1) + \cdots + \frac{1}{N}x(N-1)\right]$$
$$- \frac{N-1}{N}x(T+\alpha-1) - \cdots - \frac{1}{N}x[T+\alpha-(N-1)]$$
$$+ \left[b + \frac{N}{N+1}\right]\left\{x(T) + \frac{N-1}{N}x(T-1) + \cdots + \frac{1}{N}x[T-(N-1)]\right\} \tag{11}$$

where

$$b = \frac{1}{N+1}\left[\frac{2N - T - (N+1)\alpha}{2N-T}\right]$$

Other aspects of this work are presently being considered. Conditions on an arbitrary continuous PL function which ensure that it is a correlation function have been studied. Random processes with PL correlation functions have representations in terms of the Wiener process that are closely related to the moving average representations of arbitrary stationary random processes.

A principal aim of this work is the development of an approximate theory of prediction in the following sense: Suppose that the correlation function of the process to be predicted is approximated with a PL correlation function, and the predictor for this approximate correlation function is used for the process. It is desired to relate the degradation in prediction error to some measure of the closeness of the approximate correlation function to the true. This problem is currently under investigation.

The ideas used in this article are easily extended to include the interpolation and signal extraction problem.

G. Information Processing: Least-Squares Estimates From Likelihood Ratios, *T. Kailath*[6]

1. Introduction

In another article,[7] it was shown that the likelihood ratio for the detection of a random signal in additive white-gaussian noise could be expressed in terms of the causal least-squares estimate of the signal. This article demonstrates that, conversely, the causal least-squares estimate can be obtained from the likelihood ratio with the case of a narrowband random-phase signal worked out as an example. A related but different result in the discrete-time case is also discussed.

2. A Description of the Hypotheses

We first describe the major result of the article listed as Footnote 7. Consider the two hypotheses (H_0 and H_1)

$$\left.\begin{aligned} H_1 &: \dot{x}(t) = z(t) + \dot{w}(t) \\ H_0 &: \dot{x}(t) = \dot{w}(t) \end{aligned}\right\}, \qquad 0 \leq t \leq T \tag{1}$$

where $\dot{w}(\cdot)$ is zero-mean white-gaussian noise with covariance function

$$\overline{\dot{w}(t)\,\dot{w}(s)} = \delta(t-s) \tag{2}$$

[In other words, $w(\cdot)$ is the so-called *Wiener process*, i.e., a gaussian process with zero-mean and covariance function $\min(t,s)$.] The signal process $z(\cdot)$ is a random process (not necessarily Gaussian) that satisfies

$$\int_0^T \overline{z^2(t)}\,dt < \infty \tag{3}$$

and $z(t)$ is independent of $\dot{w}(s)$ for $s > t$. Then the likelihood ratio can be written (see Footnote 7)

$$L(T) = \exp\left[\int\!\!\!\!\!\!-_0^T \hat{z}(t|t)\,dx(t) - \frac{1}{2}\int_0^T \hat{z}^2(t|t)\,dt\right] \tag{4}$$

where $\hat{z}(t|t)$ is the least-squares estimate of $z(t)$ given $x(\tau), 0 \leq \tau \leq t$, and assuming H_1 is true, and $\int\!\!\!\!\!\!-$ denotes a special kind of stochastic integral known as the Itô integral.[8] The Itô integral has some special properties

[6]Consultant, Electrical Engineering Dept., Stanford University.

[7]Kailath, T., "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," to appear in *IEEE Trans. Inform. Theory*, 1969.

[8]Esposito, R., "On a Relation Between Detection and Estimation in Decision Theory," to appear in *Inform. Contr.*, 1968.

that differ from those of ordinary integrals. The major property is the Itô differential rule which is briefly described in *Subsection 3*. In fact, the major result of this article is that if the likelihood ratio $L(t)$ is known, then $\hat{z}(t|t)$ can be obtained from it by the formula

$$\hat{z}(t|t) = \frac{dL(t)}{L(t)\,dx(t)} \tag{5}$$

where $d(\ )$ denotes the Itô differential.

3. The Itô Differential Rule

The description of the Itô differential rule (and the Itô integral) presented here must of necessity be brief. Detailed discussions can be found in Ref. 1.

Let $b(t)$ be a stochastic process that depends, at most, on the past and present values $\{w(s), 0 \leq s \leq t\}$ of a Wiener process but is statistically independent of future values $\{w(s), s > t\}$. Such a process will be called *admissible*.

Suppose that the variance of $b(t)$ obeys

$$\int_0^T \overline{b^2(t)}\,dt < \infty$$

Then Itô showed that integrals of the form (where the bar will henceforth be used to denote Itô integrals)

$$\oint_0^t b(s)\,dw(s)$$

can be defined so that they are also admissible and continuous (with probability one). The Itô integral has certain special properties that derive basically from the fact (the so-called Lévy property) that the increments $dw(t)$ of a Wiener process are of the order of $(dt)^{1/2}$ and not $0(dt)$ as they would be for a smoother random process. This means that second-order terms $(dw)^2$ cannot be neglected in the Itô stochastic calculus. This will be made readily apparent by the Itô processes presented in *Subsection 4*. Before presenting these processes, however, the following additional definition is required.

It will be said that $\{x(t), 0 \leq t \leq T\}$ is an *Itô process* if it can be written in the form

$$x(t) = \int_0^t a(s)\,ds + \oint_0^t b(s)\,dw(s), \qquad 0 \leq t \leq T \tag{6}$$

where $a(\cdot)$ and $b(\cdot)$ are admissible processes and

$$\int_0^t |\overline{a(s)}|\,ds < \infty, \qquad \int_0^T |\overline{b(s)}|^2\,ds < \infty \tag{7}$$

It is usual to write Eq. (6) symbolically in differential form as

$$dx(t) = a(t)\,dt + b(t)\,dw(t) \tag{8}$$

Note that in this notation, $d(\ )$ and $\oint(\ )$ are inverse operations as in the usual calculus.

4. The Itô Processes

Suppose $f(x,t)$ is a function of two variables with continuous second-order partial derivatives in x and t, which will be denoted by f_t, f_x, f_{xx}, etc. Then Itô's differential rule states that $f[x(t),t]$ will also be an integral process defined by

$$df(x,t) = f_t(x,t)\,dt + f_x(x,t)\,dx + \frac{1}{2} f_{xx}(x,t)\,b^2(t)\,dt \tag{9}$$

$$= \left[f_t(x,t) + a(t) f_x(x,t) + \frac{1}{2} f_{xx}(x,t)\,b^2(t) \right] dt + b(t)\,dw(t) \tag{10}$$

The integrated form

$$f[x(t),t] - f[x(0),0] = \int_0^t \left[f_t + a f_x + \frac{1}{2} b^2 f_{xx} \right] dt + \oint_0^t f_x\,dw \tag{11}$$

is also sometimes useful. Equation (11) can be heuristically obtained by a formal Taylor expansion of $f(x + dx, t + dt)$ and use of the symbolic relations

$$dt\,dw = 0 = dx\,dw, \qquad (dw)^2 = dt$$

Note that due to the Lévy property of $w(\cdot)$, Eqs. (9)–(11) are different from the ordinary formula for the differential in that they contain the second-order term

$$\tfrac{1}{2} f_{xx}(x,t)\,b^2(t)\,dt$$

Of course, if $b(t) \equiv 0$, then the Itô formula and the ordinary formula coincide.

It will sometimes be necessary to consider the vector Itô processes

$$d\mathbf{x}(t) = \mathbf{a}(t)\,dt + \mathbf{B}(t)\,d\mathbf{w}(t) \tag{12}$$

where **a**, **w**, and **x** are vectors and **B** is a matrix. Let $f(\mathbf{x}, t)$ be a scalar function of $\mathbf{x}(t)$ and t, while $\mathbf{f}_x$ denotes the vector of first-partial derivatives and $\mathbf{f}_{xx}$ the matrix of second-partials. Then the Itô rule is

$$df(\mathbf{x},t) = f_t(\mathbf{x},t)\,dt + \mathbf{f}_x'(x,t)\,[\mathbf{a}(t)\,dt + \mathbf{B}(t)\,d\mathbf{w}(t)] + \frac{1}{2}\,tr\,[\mathbf{B}'(t)\cdot\mathbf{f}_{xx}(x,t)\,\mathbf{B}(t)] \tag{13}$$

where the prime denotes transpose.

The following two examples are given to illustrate the use of the Itô rule.

Example 1. Let

$$f(x,t) = x^2 \tag{14}$$

and

$$dx(t) = a(t)\,dt + b(t)\,dw(t) \tag{15}$$

Then

$$dx^2 = x\,dx + \frac{1}{2}\,b^2(t)\,dt \tag{16}$$

In particular, when $a(\cdot) \equiv 0$ and $b(\cdot) \equiv 1$, one has

$$\frac{1}{2}\,dw^2 = w\,dw + \frac{1}{2}\,dt \tag{17}$$

or, equivalently

$$\int_0^T w\,dw = \frac{1}{2}\int_0^T dw^2 - \frac{1}{2}\int_0^T dt = \frac{w^2(T)}{2} - \frac{T}{2} \tag{18}$$

Note that for ordinary integrals, we would not have the term $(-T/2)$.

Example 2. Let

$$f(L,t) = \ell n\,L(t) \tag{19}$$

where $dL(t) = a(t)\,dt + b(t)\,dw(t)$.

Then

$$d\,\ell n\,L(t) = \frac{dL(t)}{L(t)} - \frac{1}{2}\,\frac{1}{L^2(t)}\,b^2(t)\,dt \tag{20}$$

or, equivalently

$$\int_1^T \frac{dL(t)}{L(t)} = \ell n\,L(T) + \frac{1}{2}\int \frac{b^2(t)}{L^2(t)}\,dt \tag{21}$$

which again clearly demonstrates the difference from the usual integration rules.

5. The Estimation Formula

If, by some means, we have been able to directly determine the likelihood ratio $L(T)$ for the problem given as Eq. (1), then the formula given as Eq. (4) suggests that we should be able to determine $\hat{z}(t|t)$ from it. This is true, because direct application of the Itô differential rule to Eq. (4) yields

$$dL(t) = L(t)\cdot\hat{z}(t|t)\cdot dx(t) \tag{22}$$

Therefore

$$\hat{z}(t|t) = \frac{dL(t)}{L(t)}\cdot\frac{1}{dx(t)} \tag{23}$$

where $dL(t)$ and $dx(t)$ are the Itô differentials of $L(t)$ and $x(t)$, respectively. Note again [cf Eq. (21)] that, in the Itô calculus, $dL(t)/L(t)$ is not equal to $d\,\ell n\,L(t)$.

6. The Narrowband Random-Phase Signal Case

The formula given as Eq. (23) can be applied in any problem where $L(t)$ is easily determined. A rather trivial case is where $z(t) = \alpha m(t)$, with $m(t)$ a known signal but α a gaussian random variable. Here we shall treat the somewhat less obvious case of a narrowband signal of random phase. The likelihood ratio for the problem

$$H_1: \left\{\begin{array}{l} \dot{x}(\tau) = A(t)\cos(\omega_0 t + \theta) + \dot{w}(t) \\ P(\theta) = \dfrac{1}{2\pi} \end{array}\right\}, \quad 0 \leq \theta \leq 2\pi$$
$$H_0: \dot{x}(\tau) = \dot{w}(t) \tag{24}$$

is well known (Ref. 2) to be

$$L(t) = I_0[V(t)]\exp\left[-\frac{1}{4}\int_0^T A^2(\tau)\,d\tau\right] \tag{25}$$

where $I_0(\cdot)$ is the modified Bessel function and

$$V^2(t) = V_c^2(t) + V_s^2(t) \tag{26}$$

$$\left.\begin{array}{l} V_c(t) = \displaystyle\int_0^t A(\tau)\cos\omega_0\tau\,dx(\tau) \\ V_s(t) = \displaystyle\int_0^t A(\tau)\sin\omega_0\tau\,dx(\tau) \end{array}\right\} \tag{27}$$

To apply Eq. (23), first calculate [cf Eq. (21)]

$$\frac{dL(t)}{L(t)} = \frac{dI_0[V(t)]}{I_0[V(t)]} - \frac{1}{4}A^2(t)\,dt \tag{28}$$

Furthermore, after some algebra, and some use of the narrowband assumption, we obtain, say

$$\begin{aligned} dV(t) &= \frac{V_c(t)\,dV_c(t) + V_s(t)\,dV_s(t)}{V(t)} + \frac{1}{4}\frac{A^2(t)}{V(t)}dt \\ &= A(t)\cos[\omega_0 t + \phi(t)]\,dx(t) + \frac{1}{4}\frac{A^2(t)}{V(t)}dt \end{aligned} \tag{29}$$

where

$$\phi(t) = -\tan^{-1}\left[\frac{V_s(t)}{V_c(t)}\right] \tag{30}$$

Then

$$[dV(t)]^2 = \frac{1}{2}A^2(t)\,dt + 0(dt) \tag{31}$$

and

$$dI_0[V(t)] = \frac{\partial I_0}{\partial V}dV + \frac{1}{2}\times\frac{\partial^2 I_0}{\partial V^2}\times\frac{1}{2}A^2(t)\,dt \tag{32}$$

$$\begin{aligned} &= I_1(V)\times A(t)\times\cos[\omega_0 t + \phi(t)]\,dx \\ &\quad + \frac{1}{4}\frac{A^2(t)}{V(t)}\times\left[I_1(V) + V\frac{\partial I_1}{\partial V}\right]dt \end{aligned} \tag{33}$$

Finally, by using the Bessel function identity

$$I_1(V) + V\frac{\partial I_1}{\partial V} = VI_0(V) \tag{34}$$

and combining Eqs. (33) and (28), we obtain

$$\hat{z}(t|t) = \frac{dL(t)}{L(t)}\cdot\frac{1}{dx(t)} = \frac{I_1(V)}{I_0(V)}A(t)\cos[\omega_0 t + \phi(t)] \tag{35}$$

This formula can be directly verified.

7. The Discrete-time Case

A discussion of the likelihood ratio formula, Eq. (1), with R. Esposito (Footnote 8) led him to develop a relation in the *discrete-time* case between the likelihood ratio and the *noncausal* estimate of a random signal in gaussian noise. Thus consider the hypotheses

$$H_1: \mathbf{x} = \mathbf{z} + \mathbf{n} \qquad H_0: \mathbf{x} = \mathbf{n} \tag{36}$$

where $\mathbf{n}$ is a gaussian vector with

$$\bar{\mathbf{n}} = 0 \qquad \overline{\mathbf{n}\mathbf{n}'} = \mathbf{I} \tag{37}$$

and $\mathbf{z}$ is an independent random vector with, say, a density function $p_z(\cdot)$. Then the likelihood ratio is readily seen to be

$$\Lambda(\mathbf{x}) = \int \exp\left\{\mathbf{u}'\mathbf{x} - \frac{1}{2}\mathbf{u}'\mathbf{u}\right\} p_z(\mathbf{u})\,d\mathbf{u} \tag{38}$$

The least-squares estimate of $\mathbf{z}$, given the whole vector $\mathbf{z} + \mathbf{n} = \mathbf{x}$, is well known to be (the conditional mean)

$$\hat{\mathbf{z}} = \frac{\int \mathbf{u}\exp\left\{\mathbf{u}'\mathbf{x} - \frac{1}{2}\mathbf{u}'\mathbf{u}\right\} p_z(\mathbf{u})\,d\mathbf{u}}{\int \exp\left\{\mathbf{u}'\mathbf{x} - \frac{1}{2}\mathbf{u}'\mathbf{u}\right\} p_z(\mathbf{u})\,d\mathbf{u}} \tag{39}$$

The expressions in Eqs. (38) and (39) seem closely related; in fact, by direct differentiation of $\Lambda(\mathbf{x})$ with respect to $\mathbf{x}$ we have

$$\nabla_x \Lambda(x) = \int \mathbf{u}\exp\left\{\mathbf{u}'\mathbf{x} - \frac{1}{2}\mathbf{u}'\mathbf{u}\right\} p_z(\mathbf{u})\,d\mathbf{u} = \hat{\mathbf{z}}\,\Lambda(x) \tag{40}$$

so that

$$\hat{\mathbf{z}} = \frac{\nabla_x \Lambda(\mathbf{x})}{\Lambda(\mathbf{x})} = \nabla_x \ln \Lambda(x) \tag{41}$$

This is Esposito's relation (which he derived somewhat less directly). There is some similarity between Eq. (41) and our continuous-time formula, Eq. (5). In fact, Eq. (4) can be derived in a similar way. However, note that in discrete-time there is no Itô rule and $\nabla_x L/L$ can be written $\nabla_x \ln L$. The important difference, however, is that there is no discrete-time formula for the likelihood ratio corresponding to our general continuous-time formula, Eq. (4). All that can be concluded from Eq. (41) is that

$$L(\mathbf{x}) = \exp\left[\int \hat{\mathbf{z}}'\,d\mathbf{x} + \text{constant}\right] \tag{42}$$

Equation (42) uses noncausal estimates and is not as explicit as Eq. (4); moreover, unless an analytical expression is available for $\hat{\mathbf{z}}$ so that the integral in Eq. (42) can be evaluated analytically, it does not seem possible to implement Eq. (42) for any given observation $\mathbf{x}$. In Eq. (4), on

the other hand, the integrations are with respect to time. In addition, Eq. (4) yields a receiver structure into which suboptimum estimates for $z(\cdot)$ can easily be introduced. It will be interesting to study the relationships between the discrete-time and continuous-time analyses in more detail and, in particular, to see how to carry one over into the other.

8. Concluding Remarks

The above results can be extended to the case of colored additive gaussian noise by use of a noise-whitening filter; the results can also be easily extended to the vector case.

References

1. Skorokhod, A. V., *Studies on Random Processes* (translation). Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1965.
2. Helstrom, C. W., *Statistical Theory of Signal Detection*. Pergamon Press, Inc., New York, 1960.

H. Astrometrics: Toeplitz Matrix Inversion—The Algorithm of W. F. Trench, *S. Zohar*

1. Introduction

A matrix T whose ij element, T_{ij}, is a function of $(i-j)$, rather than of i, j separately, is called a *Toeplitz matrix*.

By way of illustration, we show here a Toeplitz matrix of order 4

$$\begin{bmatrix} \tau_0 & \tau_{-1} & \tau_{-2} & \tau_{-3} \\ \tau_1 & \tau_0 & \tau_{-1} & \tau_{-2} \\ \tau_2 & \tau_1 & \tau_0 & \tau_{-1} \\ \tau_3 & \tau_2 & \tau_1 & \tau_0 \end{bmatrix}$$

A typical example of a situation where the inverse of such a matrix is needed is the case of a linear network prescribed in terms of its sampled output correlation in response to white-noise excitation. In determining the sampled input correlation corresponding to a given sampled output correlation, we are led to the inversion of a real symmetric Toeplitz matrix. Such an application arose in planetary radar mapping.

The inversion of a non-symmetric Toeplitz matrix is called for in the case of signals known, or assumed, to have a rational spectrum. In this case, the inversion of a non-symmetric Toeplitz matrix allows us to compute all correlation coefficients from a finite number of given coefficients.

The very special structure of a Toeplitz matrix suggests that an inversion scheme exploiting this structure would yield significant savings in time and effort. As a matter of fact, several authors have approached this problem in the past. However, a really significant contribution was made in 1964 by W. F. Trench (Ref. 1). His method is both powerful and elegant. Unfortunately, his somewhat complicated derivation tends to obscure the fact that the resulting algorithm is beautifully simple. Also, Trench mainly stresses the Hermitian Toeplitz matrix and shows only the final results without proof for the more general non-Hermitian case.

The purpose of this article is to rephrase the derivations in order to make them easier to follow, thus making this powerful tool more accessible. In this method of presentation, no significant simplifications result when the derivations are limited to the Hermitian case, and so the more general non-Hermitian case is treated here from the outset. Hermitian Toeplitz matrices are handled as a simple special case in *Subsection 9*.

The inversion algorithm does not apply to all Toeplitz matrices. In order to describe the class to which it does apply, we introduce the notion of a "strongly non-singular" matrix. An arbitrary matrix is said to be strongly non-singular when, in addition to being non-singular itself, all its principal submatrices are non-singular. (Equivalently, all its principal minors are non-zero.) The inversion algorithm applies to strongly non-singular Toeplitz matrices and, hence, to positive-definite Hermitian Toeplitz matrices.

A trivial immediate result of this constraint is that the main diagonal term τ_0 is non-zero (τ_0 is the first-order principal minor). Hence, it is permissible to divide the given matrix by τ_0 to get a form with ones along the main diagonal. It is this form that will be treated here.

2. Formulation of the Problem

We start with a brief note regarding notation:

We use Greek letters for scalars, capital letters for square matrices, and small letters for column matrices. Transposition is indicated by the symbol $\sim$. The order of the matrices is indicated by a subscript. In the absence of a subscript, the order is taken to be n. Thus

$$L \equiv L_n \text{ is an } n \times n \text{ matrix}$$

$$g \equiv g_n \text{ is an } n \times 1 \text{ matrix}$$

$$\tilde{g} \equiv \tilde{g}_n \text{ is a } 1 \times n \text{ matrix}$$

$$\lambda_n \quad \text{is a scalar}$$

The initial steps of the inversion follow the well-known bordering scheme. Therefore, we represent the Toeplitz matrix L_{n+1} and its inverse, B_{n+1}, as follows

$$L_{n+1} = \begin{bmatrix} 1 & \tilde{a} \\ r & L \end{bmatrix} \tag{1}$$

$$B_{n+1} = \frac{1}{\lambda_n} \begin{bmatrix} 1 & \tilde{e} \\ g & M \end{bmatrix} \tag{2}$$

Note that L_{n+1} is fully specified in terms of the matrices a, r

$$\tilde{r} = [\rho_1 \rho_2 \cdots \rho_n]$$

$$\tilde{a} = [\rho_{-1} \rho_{-2} \cdots \rho_{-n}]$$

The representation of the inverse adopted in Eq. (2) is not general. Thus, with all the terms involved being finite, Eq. (2) certainly rules out an inverse whose (1, 1) term is zero. This, however, is consistent with the property of strong non-singularity. To see this, we apply the general formula expressing an inverse term as a ratio of determinants.

$$(B_{n+1})_{11} = \frac{|L|}{|L_{n+1}|}$$

Now, since $|L_{n+1}|$ is finite and $|L| \neq 0$, it follows that $(B_{n+1})_{11} \neq 0$ and the representation adopted in Eq. (2) is valid.

We record here, for future use, the specific meaning of the parameter λ_n that follows from this

$$\lambda_n = \frac{|L_{n+1}|}{|L_n|} \tag{2a}$$

3. Applying the Bordering Approach

Multiplying Eqs. (2) and (1), denoting the identity matrix by I, and a zero-column matrix by $\mathcal{O}$, we obtain

$$\begin{bmatrix} 1 & \tilde{\mathcal{O}} \\ \mathcal{O} & I \end{bmatrix} = \frac{1}{\lambda_n} \begin{bmatrix} 1 + \tilde{e}r & \tilde{a} + \tilde{e}L \\ g + Mr & g\tilde{a} + ML \end{bmatrix}$$

The first column yields

$$\lambda_n = 1 + \tilde{e}r \tag{3}$$

$$g = -Mr \tag{4}$$

In deriving the results of the two remaining terms, we multiply both sides by $B (\equiv B_n \equiv L^{-1})$ and arrive at

$$\tilde{e} = -\tilde{a}B \tag{5}$$

$$M = \lambda_n B - g\tilde{a}B = \lambda_n B + g\tilde{e} \tag{6}$$

Thus, Eq. (2) now takes the form

$$\lambda_n B_{n+1} = \begin{bmatrix} 1 & \tilde{e} \\ g & \lambda_n B + g\tilde{e} \end{bmatrix} \tag{7}$$

Equations (3) to (6) facilitate the expression of all the elements of B_{n+1} as functions of B_n. This is the basic idea of the bordering algorithm. Here, however, we can dispense with this cumbersome approach because the inverse of a Toeplitz matrix is completely prescribed in terms of its first (or last) row and column.

The proof of this property is facilitated by the introduction of the notions of persymmetry and the exchange matrix E taken up in *Subsection 4.*

4. Persymmetry and the Exchange Matrix

A square matrix is said to be persymmetric when it has symmetry about its cross diagonal (the diagonal extending from the upper-right corner to the lower-left corner). If $P (\equiv P_n)$ is persymmetric, then

$$P_{ij} = P_{n+1-j,\, n+1-i}$$

Operations with these matrices are facilitated by the use of the exchange matrix E, which is defined as a square matrix with units along the cross diagonal and zeros elsewhere. Premultiplying a matrix by E exchanges its elements which are located symmetrically about a bisecting horizontal line. Postmultiplying does the same with respect to a bisecting vertical line.

To see the connection of E with the concept of persymmetry, let us examine the expression $E\,\tilde{A}\,E$, where A is an arbitrary $(n \times n)$ matrix. Visualizing the effect of the three indicated operations on a single element of A, we find that the overall effect is to exchange elements that are located symmetrically with respect to the cross diagonal. Hence, we can rephrase the definition of a persymmetric matrix as a matrix P satisfying

$$E\,\tilde{P}\,E = P$$

From this formulation, it is easy to see that the inverse of a persymmetric matrix is persymmetric. Starting with

$$\tilde{P}^{-1}\tilde{P} = I \tag{8}$$

and noting that $EE = I$, we can rephrase Eq. (8) as

$$(E\tilde{P}^{-1}E)(E\tilde{P}E) = I$$

But $E\tilde{P}E = P$. Hence

$$E\tilde{P}^{-1}E = P^{-1}$$

and P^{-1} is persymmetric as was to be shown.

5. The Self-Regeneration of the Toeplitz Inverse

We now note that a Toeplitz matrix is a special case of a persymmetric matrix. Hence, we conclude that the inverse of a Toeplitz matrix is persymmetric. Since both L and L_{n+1} appearing in Eq. (1) are Toeplitz matrices, it follows that their inverses, B_{n+1} and B, are persymmetric. This is the basic property needed to establish the algorithm for generating elements of B_{n+1} from its first row and column.

To do this, we apply the condition of persymmetry to Eq. (7)

$$\lambda_n B_{n+1} = E_{n+1}\begin{bmatrix} 1 & \tilde{g} \\ e & \lambda_n \tilde{B}_n + e\tilde{g} \end{bmatrix} E_{n+1} \tag{9}$$

In expanding expressions such as these, it is convenient to have a special symbol for the product of E by a column matrix. We use the notation $Eg = \hat{g}$ with similar expressions for the other column matrices. (Note that the elements of $\hat{g}$ are just the elements of g in reversed order). With this notation, and the fact that B is persymmetric, Eq. (9) leads to

$$\lambda_n B_{n+1} = \begin{bmatrix} 1 & \tilde{e} \\ g & \lambda_n B + g\tilde{e} \end{bmatrix} = \begin{bmatrix} \lambda_n B + \hat{e}\tilde{\hat{g}} & \hat{e} \\ \hat{\tilde{g}} & 1 \end{bmatrix} \tag{10}$$

We now express an element of $\lambda_n B_{n+1}$ in terms of the middle matrix

$$(\lambda_n B_{n+1})_{i+1,\,j+1} = \lambda_n B_{ij} + (g\tilde{e})_{ij} \qquad 1 \leq i, j \leq n \tag{11}$$

To eliminate B_{ij}, we now express an element of $\lambda_n B_{n+1}$ in terms of the right hand matrix [in Eq. (10)]

$$\lambda_n (B_{n+1})_{ij} = \lambda_n B_{ij} + (\hat{e}\tilde{\hat{g}})_{ij} \qquad 1 \leq i, j \leq n \tag{12}$$

combining Eqs. (11) and (12), we obtain

$$\lambda_n (B_{n+1})_{i+1,\,j+1} = \lambda_n (B_{n+1})_{ij} + (g\tilde{e} - \hat{e}\tilde{\hat{g}})_{ij} \qquad 1 \leq i, j \leq n \tag{13}$$

Thus, given an element of B_{n+1}, we can generate all the remaining elements along the same diagonal. The entities needed for this are λ_n, g_n, e_n. But these are, essentially, the elements of the first row and column of B_{n+1} [see Eq. (2)] and as such, they supply the initial values for the recursion given in Eq. (13).

The dependence of the proof of this result on the fact that B_{n+1} is the inverse of a Toeplitz matrix is subtle and merits some elaboration. Obviously, the crucial step is embodied in Eq. (10). Now, while it is true that any persymmetric matrix B_{n+1} (or, for that matter, any strongly non-singular matrix) can be represented in the specific form shown in Eq. (7), the resulting B matrix will, in general, not be persymmetric. It is the persymmetry of B that is essential to the derivation of Eq. (10).

It should be noted that only about half of the elements of B_{n+1} have to be computed in this way, the remainder being given by persymmetry. We point out the computational simplicity of Eq. (13) which becomes evident when the function of g, e appearing there is written down explicitly. Thus

$$(g\tilde{e} - \hat{e}\tilde{\hat{g}})_{ij} = g_{i1}e_{j1} - e_{n+1-i,1}g_{n+1-j,1} \tag{14}$$

6. The Recursion Relations

We now proceed to establish recursion relations for the quantities λ_i, e_i, g_i $(1 \leq i \leq n)$. In view of the self-regeneration property of B_{n+1}, this is, essentially, all that is needed to obtain this matrix.

In deriving these relations, we use Eqs. (5), (7), and (10) and rely on the fact that they are valid for matrices of any order rather than the basic order n indicated.

Starting with the e_i recursion, we apply Eq. (5) to e_{i+1} and obtain

$$e_{i+1} = -\tilde{B}_{i+1}a_{i+1} \qquad 1 \leq i < n$$

where a_{i+1} consists of the first $i+1$ elements of a.

Since B_{i+1} is persymmetric, this can be written as

$$e_{i+1} = -(E_{i+1}B_{i+1}E_{i+1})a_{i+1} = -E_{i+1}B_{i+1}\hat{a}_{i+1}$$

Writing this out explicitly using Eq. (7) we obtain

$$e_{i+1} = -\frac{1}{\lambda_i} E_{i+1} \begin{bmatrix} 1 & \tilde{e}_i \\ g_i & \lambda_i B_i + g_i \tilde{e}_i \end{bmatrix} \begin{bmatrix} \rho_{-(i+1)} \\ \hat{a}_i \end{bmatrix} \qquad 1 \leq i < n \tag{15}$$

Note that we have expressed $\hat{a}_{i+1}$ as $\hat{a}_i$ augmented by a single term. Carrying out the indicated multiplications, and again applying Eq. (5), we arrive at the desired recursion relationship

$$e_{i+1} = \begin{bmatrix} e_i \\ 0 \end{bmatrix} - \frac{\rho_{-(i+1)} + \tilde{e}_i \hat{a}_i}{\lambda_i} \begin{bmatrix} \hat{g}_i \\ 1 \end{bmatrix} \qquad 1 \leq i < n \tag{16}$$

We again note the relative simplicity of the computations. The recursion relation for g can be obtained directly from Eq. (16) by noting the symmetry of the basic equations. Thus, we obtain

$$g_{i+1} = \begin{bmatrix} g_i \\ 0 \end{bmatrix} - \frac{\rho_{i+1} + \tilde{g}_i \hat{r}_i}{\lambda_i} \begin{bmatrix} \hat{e}_i \\ 1 \end{bmatrix} \qquad 1 \leq i < n \tag{17}$$

Premultiplying both sides by E_{i+1}, we get the final form used in the algorithm

$$\hat{g}_{i+1} = \begin{bmatrix} 0 \\ \hat{g}_i \end{bmatrix} - \frac{\rho_{i+1} + \tilde{r}_i \hat{g}_i}{\lambda_i} \begin{bmatrix} 1 \\ e_i \end{bmatrix} \qquad 1 \leq i < n \tag{18}$$

Here, we have used the fact that $\tilde{g}_i \hat{r}_i = \tilde{r}_i \hat{g}_i$ so that in Eqs. (16) and (18), the g matrices appear consistently in their "exchanged" form ($\hat{g}$).

The recursion for λ is somewhat simpler. Writing Eq. (10) explicitly for the bottom-right element, and replacing n by i, we obtain

$$\lambda_i (B_{i+1})_{i+1,\, i+1} = \lambda_i (B_i)_{ii} + (g_i)_{i1} (e_i)_{i1} = 1 \tag{19}$$

Replacing i by $i+1$ in the middle term gives us

$$\lambda_{i+1} (B_{i+1})_{i+1,\, i+1} + (g_{i+1})_{i+1,\, 1} (e_{i+1})_{i+1,\, 1} = 1 \tag{20}$$

Eliminating B_{i+1} from Eqs. (19) and (20), we arrive at the recursion relationship for λ

$$\lambda_{i+1} = \lambda_i [1 - (g_{i+1})_{i+1,\, 1} (e_{i+1})_{i+1,\, 1}] \qquad 1 \leq i < n \tag{21}$$

Using Eqs. (16) and (18), this can be expressed in terms of e_i, g_i as

$$\lambda_{i+1} = \lambda_i - \frac{(\rho_{i+1} + \tilde{r}_i \hat{g}_i)(\rho_{-(i+1)} + \tilde{e}_i \hat{a}_i)}{\lambda_i} \qquad 1 \leq i < n \tag{22}$$

Equations (16), (18), and (22) require division by λ_i. This is always possible because the strong non-singularity of L_{n+1} guarantees $\lambda_i \neq 0$ $(1 \leq i \leq n)$. To see this, we recall Eq. (2a), repeating it here for the index i.

$$\lambda_i = \frac{|L_{i+1}|}{|L_i|} \qquad 1 \leq i \leq n \tag{23}$$

With $|L_i|$ finite and $|L_{i+1}| \neq 0$, we must have $\lambda_i \neq 0$. This is the motivation for imposing the constraint of strong non-singularity.

We note that in the case of Hermitian matrices, this constraint is much weaker than the positive-definiteness constraint originally imposed by W. F. Trench[9] (Ref. 1).

Returning now to the recursion relations, we note that the initial values can be obtained rather trivially by considering $i = 1$

$$L_2 = \begin{bmatrix} 1 & \rho_{-1} \\ \rho_1 & 1 \end{bmatrix}$$

$$B_2 = \frac{1}{1 - \rho_{-1}\rho_1} \begin{bmatrix} 1 & -\rho_{-1} \\ -\rho_1 & 1 \end{bmatrix}$$

Hence

$$\lambda_1 = 1 - \rho_{-1}\rho_1$$

$$e_1 = -\rho_{-1}$$

$$g_1 = -\rho_1$$

This completes the derivation of the inversion algorithm. The overall scheme is summarized in *Subsection 7.*

[9]In discussing the constraint for the non-Hermitian algorithm, Trench does mention the correct weaker constraint. However, in his discussion of the Hermitian algorithm, he specifically imposes the stronger constraint of positive-definiteness.

7. The Trench Algorithm (Non-Hermitian Case)

Problem formulation.

$$L_{n+1} = \begin{bmatrix} 1 & \tilde{a} \\ r & L_n \end{bmatrix}$$

$$(a_n)_{i1} = \rho_{-i}$$

$$(r_n)_{i1} = \rho_i$$

$$B_{n+1} = L_{n+1}^{-1}$$

Initial values for recursion.

$$\lambda_1 = 1 - \rho_{-1}\rho_1$$

$$e_1 = -\rho_{-1}$$

$$g_1 = -\rho_1$$

Recursion of $\lambda, g, e\ (1 \leq i < n)$.

$$\eta_i = -(\rho_{-(i+1)} + \tilde{e}_i \hat{a}_i)$$

$$\gamma_i = -(\rho_{i+1} + \tilde{r}_i \hat{g}_i)$$

$$e_{i+1} = \begin{bmatrix} e_i + \frac{\eta_i}{\lambda_i}\hat{g}_i \\ \\ \frac{\eta_i}{\lambda_i} \end{bmatrix}, \qquad \hat{g}_{i+1} = \begin{bmatrix} \frac{\gamma_i}{\lambda_i} \\ \\ \hat{g}_i + \frac{\gamma_i}{\lambda_i} e_i \end{bmatrix}$$

$$\lambda_{i+1} = \lambda_i - \frac{\eta_i \gamma_i}{\lambda_i}$$

Evaluation of B_{n+1}.

$$(B_{n+1})_{11} = \frac{1}{\lambda_n}$$

$$(B_{n+1})_{1,j+1} = \frac{(e_n)_{j1}}{\lambda_n} \qquad 1 \leq j \leq n$$

$$(B_{n+1})_{i+1,1} = \frac{(g_n)_{i1}}{\lambda_n} \qquad 1 \leq i \leq n$$

$$(B_{n+1})_{i+1,j+1} = (B_{n+1})_{ij} + \frac{1}{\lambda_n}(g\tilde{e} - \hat{e}\tilde{g})_{ij} \qquad 1 \leq i, j < n$$

$$(B_{n+1})_{ij} = (B_{n+1})_{n+2-j,\,n+2-i} \quad \text{(persymmetry)}$$

8. Computer Implementation

The most significant aspect of the above algorithm is that, although we are dealing with the inversion of a matrix, most of the computations involve one-dimensional arrays. Thus, all the information prescribed by L_{n+1} is contained in its first row and column. We arrange these into a single row matrix as

$$[\tilde{a}\ 1\ \tilde{r}]$$

and store it in the computer as the following one-dimensional array

$\rho_{-n}, \cdots, \rho_{-3}, \rho_{-2}, \rho_{-1}$	1	$\rho_1, \rho_2, \rho_3, \cdots, \rho_n$

We refer to this as the R array.

As we have seen, B_{n+1} is completely prescribed in terms of e, g, λ_n which we arrange in a single-row matrix

$$[\tilde{g}\ \lambda_n\ \tilde{e}]$$

The corresponding computer array is denoted X and has the following structure

$(g_n)_{n1}, \cdots, (g_n)_{11}$	λ_n	$(e_n)_{11}, \cdots, (e_n)_{n1}$

The main computational effort is in obtaining array X from the given array R. We start the process using the three central elements of R to assign (transient) values to the three central elements of X, namely $(g_1)_{11}$, λ_1, $(e_1)_{11}$. In the next cycle, two more elements of R, (ρ_{-2}, ρ_2), are used to fill in two more cells of X. This process of spreading from the center to the edges continues until all the cells of X are filled up.

The details of the algorithm are illustrated in Fig. 6. The third array Y shown here is a temporary storage array that can be eliminated with a saving of both time and storage. It is adopted here only in the interest of a simpler representation of the algorithm. All three arrays are of size $(2n + 1)$ and are assumed to have an index range $(-n, n)$.

Figure 6 describes the situation immediately after e_3, g_3, λ_3 have been determined. ρ_4 (plus the products involving ρ_1, ρ_2, ρ_3) yields γ_3. Similarly, ρ_{-4} (plus the products involving $\rho_{-3}, \rho_{-2}, \rho_{-1}$) yields η_3. This permits us to enter the two new values of X

$$X(-4) \equiv (g_4)_{41} = \frac{\gamma_3}{\lambda_3}$$

$$X(4) \equiv (e_4)_{41} = \frac{\eta_3}{\lambda_3}$$

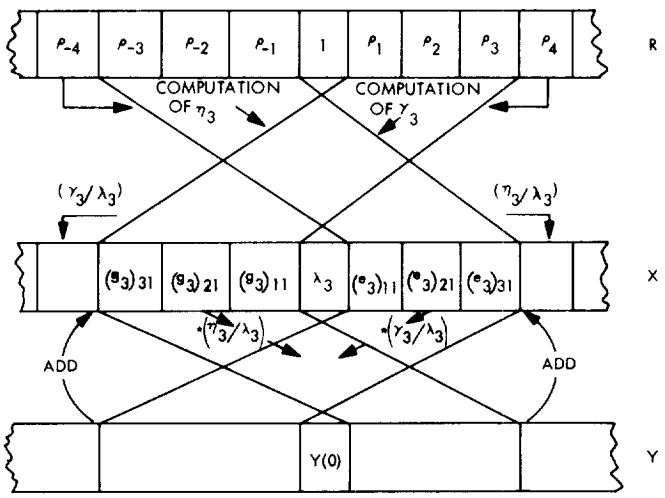


Fig. 6. Algorithm for the general case

The updating of the remaining seven terms of X is now accomplished using the temporary storage array Y. First, we place the indicated multiples of e_3 and g_3 in Y. We also set

$$Y(0) = -\frac{\eta_3\gamma_3}{\lambda_3}$$

Now, upon adding the seven central elements of Y to the corresponding elements of X, we complete the cycle and are ready to start all over with the next cycle.

9. The Trench Algorithm (Hermitian Case)

When L_{n+1} is Hermitian, we have $a_i = r_i^*$ (* denotes complex conjugation) and, consequently, $e_i = g_i^*$. Substituting these in the formulae of *Subsection 7*, we arrive at the following summary for this important special case.

Problem formulation.

$$L_{n+1} = \begin{bmatrix} 1 & \tilde{r}^* \\ r & L_n \end{bmatrix}$$

$$(r_n)_{i1} = \rho_i$$

$$B_{n+1} = L_{n+1}^{-1}$$

Initial values for recursion.

$$\lambda_1 = 1 - |\rho_1|^2$$

$$g_1 = -\rho_1$$

Recursion of $\lambda, g (1 \leq i < n)$.

$$\gamma_i = -(\rho_{i+1} + \tilde{r}_i \hat{g}_i)$$

$$\tilde{g}_{i+1} = \begin{bmatrix} \dfrac{\gamma_i}{\lambda_i} \\ \\ \hat{g}_i + \dfrac{\gamma_i}{\lambda_i} g_i^* \end{bmatrix}$$

$$\lambda_{i+1} = \lambda_i - \frac{|\gamma_i|^2}{\lambda_i}$$

Evaluation of B_{n+1}.

$$(B_{n+1})_{11} = \frac{1}{\lambda_n}$$

$$(B_{n+1})_{i+1,1} = \frac{(g_n)_{i1}}{\lambda_n} \qquad 1 \leq i \leq n$$

$$(B_{n+1})_{i+1,j+1} = (B_{n+1})_{ij} + \frac{1}{\lambda_n}(g\tilde{g}^* - \hat{g}^*\hat{\tilde{g}})_{ij} \qquad 1 \leq i,j < n$$

$$(B_{n+1})_{ij} = (B_{n+1})_{n+2-j,\,n+2-i} \qquad \text{(persymmetry)}$$

$$(B_{n+1})_{ij} = (B_{n+1})_{ji}^* \qquad \text{(Hermitian)}$$

We now briefly consider the computer implementation in this case. The situation corresponding to Fig. 6 of the general case is shown in Fig. 7. As a special case derived from Fig. 6 this is mostly self-explanatory. Note

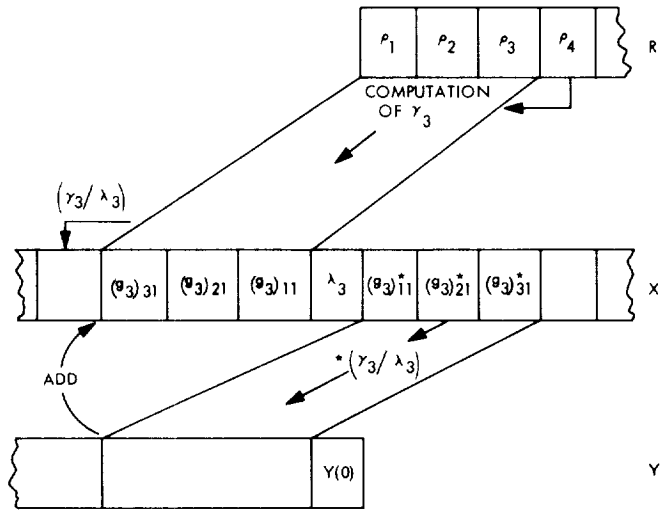


Fig. 7. Algorithm for the Hermitian case

that $Y(0)$ is now given by

$$Y(0) = -\frac{|\gamma_3|^2}{\lambda_3}$$

Also note that because both g_i and $\hat{g}_i^*$ are needed, array X still has the dimension $(2n + 1)$ as in the general case.

Figure 7 describes specifically the steps involved in computing the left half of array X, namely $\hat{g}_4$. The right half, g_4^* is obtained by "reflection" of the left half about $X(0)$, followed by the operation of complex conjugation.

10. Computing $|L_{n+1}|$

If in addition to computing the inverse of L_{n+1} we also want its determinant, the extra effort amounts to just n multiplications. The relevant formula is derived from Eq. (23).

$$|L_{i+1}| = \lambda_i |L_i| \qquad 1 \leq i \leq n$$

This is, in fact, a recursion relation for the computation of $|L_{n+1}|$. Recalling [Eq. (1)] that $|L_1| = 1$, we arrive at the final result

$$|L_{n+1}| = \prod_{i=1}^{n} \lambda_i \tag{24}$$

11. Concluding Remarks

As anticipated, the final algorithm represents an appreciable saving when compared to an inversion ignoring the Toeplitz structure. We wish to point out that in some applications (e.g., operations with rational spectra) the saving is even greater than so far indicated. It turns out that the first row and column of the inverse, rather than the complete inverse, are all that is needed to solve the problem at hand. Thus, all the computations involved in generating B_{n+1} from e, g, λ_n can be dispensed with.

Another point to bear in mind is that when dealing with very large Toeplitz matrices, an appreciable saving in memory requirements can be effected by using special subroutines that would perform the multiplication of B_{n+1} by a column or square matrix without ever having to assign the $(n + 1)^2$ memory cells to store B_{n+1}. These subroutines would incorporate the self-regeneration algorithm so that the storage requirements for B_{n+1} would be just the $(2n + 1)$ cells of array X.

Reference

1. Trench, W. F., "An Algorithm for the Inversion of Finite Toeplitz Matrices," *J. Soc. Indust. Appl. Math.*, Vol. 12, No. 3, pp. 515–522, Sept. 1964.

I. Data Compression Techniques: Entropy of Graphs, *E. C. Posner*

1. Introduction

This article is concerned with the definition of an ϵ-entropy concept for finite graphs, and the relationship of this concept to the problem of the efficient representation of random sources of data. The concept of a random data source is idealized as a probabilistic metric space[10] (Ref. 1). Such a space is a complete, separable metric space (X, d) together with a probability measure μ on the Borel subsets of X. The points x of X represent outcomes of experiments made according to the distribution μ. The distance $d(x, y)$ indicates the loss of fidelity if x actually occurs when y is thought to occur. It is desired to describe all but a set of probability 0 of X in such a way that two close points of X need not be distinguished. That is, given $\epsilon > 0, X$ is to be partitioned by a countable union of Borel sets of diameters of at most ϵ (i.e., by an ϵ-partition), and it is necessary to know into which set of the partition the actual outcome falls. This is the general problem of data compression.[11]

In order to determine the number of bits of information $H(\mathcal{U})$ necessary to describe into which set U_i of the ϵ-partition $\mathcal{U}$ the outcome x falls let

$$p_i = \mu(U_i) \tag{1}$$

Then

$$H(\mathcal{U}) = \sum_i p_i \log \frac{1}{p_i} \tag{2}$$

is the entropy of the ϵ-partition $\mathcal{U}$. A word of caution: If each outcome x must be described separately, the most efficient assignment scheme of binary words to sets U_i may have an average length greater than $H(\mathcal{U})$. It is possible to achieve $H(\mathcal{U})$ asymptotically only if long sequences of independent outcomes are observed and the entire sequence of sets of the partition is then described. If $H_1(\mathcal{U})$ represents the actual attainable value with the best encoding scheme on a single experiment, then (Ref. 2, p. 71)

$$H(\mathcal{U}) \leq H_1(\mathcal{U}) < H(\mathcal{U}) + \log 2$$

[10]Posner, E. C., Rodemich, E. R., and Rumsey, H., Jr., "Epsilon Entropy of Gaussian Processes" (submitted to *Ann. Math. Stat.*).

[11]Posner, E. C., and Rodemich, E. R., "Epsilon Entropy and Data Compression" (in preparation).

In this article, we are interested mainly in the case $H(\mathcal{U})$ large. Therefore, we will ignore the difference between H and H_1 and adopt H.

Since the partition $\mathcal{U}$ may not have been particularly well chosen from the point of view of minimizing $H(\mathcal{U})$, we define

$$H_\epsilon(X) = \inf_{\substack{\mathcal{U} \text{ an } \epsilon\text{-partition} \\ \text{of } X}} [H(\mathcal{U})] \tag{3}$$

as the ϵ-entropy of X; in Ref. 1, Theorem 2, it is shown that $H_\epsilon(X) = H(\mathcal{U})$ for some ϵ-partition U of X. Thus, $H_\epsilon(X)$ represents the minimum number of bits necessary to describe experimental outcomes to within fidelity ϵ, a property of the probabilistic metric space X. The ϵ-entropy $H_\epsilon(X)$ is infinite if and only if every ϵ-partition $\mathcal{U}$ of X has infinite entropy $H(\mathcal{U})$. (The separability property of the metric space X guarantees that the class of ϵ-partitions is indeed non-empty.)

There are numerous examples of probabilistic metric spaces of practical interest. For example, Ref. 1 and Footnote 10 are concerned with mean-continuous stochastic processes on the unit interval with L_2-norm as the metric, and with probability measure μ on $L_2[0,1]$ induced by the joint distributions of the process. Another important example consists of compact manifolds under a probability measure invariant under some Lie group of motions of the manifold, e.g., the n-sphere with probability measure proportional to hypersurface area (see Footnote 11). The stochastic process class of examples can be considered to give rise to problems on the efficient approximation of random functions. Similarly, the manifold examples give rise to problems of efficient packing.

2. Graphs and Their Entropy

In the remainder of this article, we shall be concerned with probabilistic metric spaces arising from finite graphs. The purpose of this article is to reduce problems arising for arbitrary probabilistic metric spaces to the same problem arising for the smaller class of spaces derived from graphs. These problems are concerned with the effect of allowing many experiments to be performed and then using an ϵ-partition of the product space, the object being to save bits by allowing storage of experiments. As mentioned earlier, little saving can be effected if the same partition is used in each factor of the product space. The question is whether the same result is true if arbitrary ϵ-partitions are allowed for the product space.

We will now proceed to define the probabilistic metric space associated with a graph. (All graphs are unoriented, finite, and have no loops or multiple vertices.) Let G be such a graph with n vertices v, and let μ be such that

$$\mu(\{v\}) = \frac{1}{n} \tag{4}$$

i.e., each vertex is equally likely. To define the metric d, let $d(v,v) = 0$ (all $v \in G$), and, in general, let $d(v,w)$ equal the length of the shortest path connecting v and w (if v and w are connected) where each edge has length 1. If

$$m = \max_{v,w \text{ connected}} d(v,w)$$

let $d(v,w) = m + 1$ if v and w are not connected.

Thus, $d(v,w) = 1$, if and only if, v and w are adjacent; $d(v,w) > 1$ if and only if v and w are unequal and non-adjacent. For the purposes of this article, any definition of $d(v,w)$ for v and w unequal and non-adjacent has the same effect as any other definition, as long as such $d(v,w)$ are defined to be greater than 1; all that matters is whether v and w are equal or adjacent, or not.

The entropy $H(G)$ of the graph G is now defined as the 1-entropy of G regarded as a probabilistic metric space. That is, $H(G)$ is the minimum of the entropy of partitions of the graph G by complete subgraphs (or "cliques"), because a set of vertices of G of a diameter of at most 1 is a complete subgraph.

There is no loss of generality in studying the 1-entropy of graphs instead of, say, the k-entropy, for k an integer greater than 1, because $H_k(G)$ is equal to $H_1(G_k)$, where G_k is the graph obtained from G by placing a new edge between every pair v, w of non-adjacent vertices of G at a distance of at most k.

We will see in *Subsection 4* that this article's questions of interest concerning the ϵ-entropy of arbitrary probabilistic metric spaces can be reduced to the corresponding question about the entropy of graphs constructed from the metric spaces. This is the main reason for studying the entropy of graphs, although the concept of entropy for graphs is interesting in its own right as it provides a framework for discussing the complexity of graphs different from other measures of complexity (Ref. 3; Ref. 4, Chap. 2; Ref. 5, Chap. 4). Also, since the general questions presented in this article have a direct application in random approximation and packing theory, *Subsection 4*

will present a possible application of graph theory to other areas of mathematics.

3. Products of Graphs and Probabilistic Metric Spaces

In this subsection, we will construct the n-fold product probabilistic metric space $X^{(n)}$ of the probabilistic metric space X. The metric $d^{(n)}$ on $X^{(n)}$ is defined as

$$d^{(n)}[(x_1, x_2, \cdots, x_n); (y_1, y_2, \cdots, y_n)] = \max_{1 \leq i \leq n} d(x_i, y_i) \tag{5}$$

This is the appropriate metric when it is required that, in a sequence of experiments, every outcome be known within ϵ. The probability measure $\mu^{(n)}$ is, of course, product measure.

If G is a graph, the probabilistic metric space $G^{(n)}$ is identified with the graph $G^{(n)}$ called the n-fold tensor product of G with itself. The tensor product $G \bigotimes H$ of two graphs has for its vertex set the cartesian product of the vertex sets of G and H; two ordered pairs in $G \bigotimes H$ are equal or adjacent in $G \bigotimes H$ if and only if corresponding component vertices are equal or adjacent in their own graphs. This agrees with the definition of $d^{(n)}$. The term "tensor product" is used because the adjacency matrix of $G \bigotimes H$ (ones are required down the diagonal in this usage of adjacency matrix) is the tensor product of the adjacency matrix of G with the adjacency matrix of H.

Let us consider the two functions $H_\epsilon[X^{(n)}]$ and $H[G^{(n)}]$. It is easy to obtain

$$H_\epsilon[X^{(n)}] \leq nH_\epsilon(X) \tag{6}$$

and so

$$H[G^{(n)}] \leq nH(G) \tag{7}$$

This is because the entropy of a "product partition" of $X^{(n)}$ (i.e., of a partition of $X^{(n)}$ consisting of cartesian products of sets in partitions of the separate factors) is equal to the sum of the entropies of the partitions in the individual factors.[12]

The entropy $H_\epsilon[X^{(n)}]$ is a measure of the number of bits necessary to describe the outcomes of independent experiments from X when the storage of n outcomes is allowed before transmission. The saving comes from not having to use product partitions by using a more efficient partition of $X^{(n)}$. This article is concerned with the savings possible in data compression by allowing such storage. This is a really fundamental question in information theory, the answer to which is yet to be found. However, it will be shown that the fractional saving achieved for graphs is the same as the fractional saving achievable for arbitrary probabilistic metric spaces. In other words, for this class of problem, it is enough to study graphs.

4. The Storage Constants

This subsection defines the storage constants $A_\lambda^{(n)}, B_\lambda^{(n)}$, $A^{(n)}$, $B^{(n)}$, $\underline{A}$, $\underline{B}$, A_λ, B_λ, A, B, for n a positive integer ($\lambda > 0$). First, define

$$A_\lambda^{(n)} = \sup_{\infty > H_\epsilon(X) > \lambda} \left\{ \frac{nH_\epsilon(X)}{H_\epsilon[X^{(n)}]} \right\} \tag{8}$$

called the n, λ storage constant. It represents the supremum of the fractional bit savings possible for spaces of ϵ-entropy more than λ when storage of n-experiments is allowed before partitioning. The n, λ graph-storage constant $B_\lambda{}^{(n)}$ is defined by

$$B_\lambda^{(n)} = \sup_{H(G) > \lambda} \left\{ \frac{nH(G)}{H[G^{(n)}]} \right\} \tag{9}$$

Both $A_\lambda^{(n)}$ and $B_\lambda^{(n)}$ are less than n. The n storage constant $A^{(n)}$ and the n-graph storage constant $B^{(n)}$ are defined by

$$\left.\begin{aligned} A^{(n)} &= \lim_{\lambda \to \infty} A_\lambda^{(n)} \\ B^{(n)} &= \lim_{\lambda \to \infty} B_\lambda^{(n)} \end{aligned}\right\} \tag{10}$$

The limits in Eq. (10) exist because $A_\lambda^{(n)}$ and $B_\lambda^{(n)}$ are decreasing in λ.

Before defining the absolute storage constants, we observe, as in Footnote 11, that

$$H_\epsilon[X^{(m+n)}] \leq H_\epsilon[X^{(m)}] + H_\epsilon[X^{(n)}] \tag{11}$$

Hence

$$\lim_{n \to \infty} \frac{1}{n} H_\epsilon[X^{(n)}] = \overline{H}_\epsilon(X) \tag{12}$$

exists (possibly infinite), and is called the absolute epsilon entropy of X. Similarly, the absolute entropy $\overline{H}(G)$ of a graph G is defined as

$$\lim_{n \to \infty} H[G^{(n)}] = \overline{H}(G) \tag{13}$$

[12]Posner, E. C., Rodemich, E. R., and Rumsey, H., Jr., "Product Entropy of Gaussian Distributions," (submitted to *Ann. Math. Statist.*)

The quantity $\overline{H}_\epsilon(X)$ is the number of bits per experiment needed to describe arbitrarily long blocks of outcomes of X to within ϵ; this is made quite precise in the article referenced in Footnote 11.

Because of Eq. (11), we see that the limits

$$\underline{A} = \lim_{n\to\infty} A^{(n)} \tag{14}$$

$$\underline{B} = \lim_{n\to\infty} B^{(n)} \tag{15}$$

exist, although they are possibly infinite. They can be called the limit storage constant and the limit graph storage constant, respectively.

The absolute λ-storage constant A_λ is then defined by

$$A_\lambda = \sup_{\infty > H_\epsilon(X) > \lambda} \left\{\frac{H_\epsilon(X)}{\overline{H}_\epsilon(X)}\right\} \tag{16}$$

and the absolute graph λ-storage constant B_λ by

$$B_\lambda = \sup_{H(G)>\lambda} \frac{H(G)}{\overline{H}(G)} \tag{17}$$

Finally, the absolute storage constant A is defined by

$$A = \lim_{\lambda\to\infty} A_\lambda \tag{18}$$

the absolute graph-storage constant B by

$$B = \lim_{\lambda\to\infty} B_\lambda \tag{19}$$

It is apparent from Eqs. (6) and (7) that all $A_\lambda^{(n)}$, $B_\lambda^{(n)}$, $A^{(n)}$, $B^{(n)}$, A_λ, B_λ, $\underline{A}$, $\underline{B}$, A, B are at least 1, and $A_\lambda^{(n)}$, $B_\lambda^{(n)}$, A_λ, B_λ can be shown by examples to be actually greater than 1; for all that is known, $\underline{A}$, $\underline{B}$, A_λ, B_λ, A, B may be infinite. However, it is easy to show that $A_\lambda^{(n)}$, $\underline{B}_\lambda^{(n)}$, $A^{(n)}$, $\underline{B}^{(n)}$ must be less than n for $n > 1$. Also, $\overline{A} \leq A$ and $\overline{B} \leq B$. We define, but do not name

$$\underline{A}_\lambda = \lim_{n\to\infty} A_\lambda^{(n)} \tag{20}$$

$$\underline{B}_\lambda = \lim_{n\to\infty} B_\lambda^{(n)} \tag{21}$$

Similar comments apply to these. We do not bother to further mention

$$\underline{\underline{A}} = \lim_{\lambda\to\infty} \underline{A}_\lambda, \qquad \underline{\underline{B}} = \lim_{\lambda\to\infty} \underline{B}_\lambda$$

The main open problem in the theory of epsilon entropy of probabilistic metric spaces is to find the absolute storage constant A. More specifically, it is desirable to know whether A is infinite, or even whether A is greater than 1. That is, it is necessary to know how much can be saved in transmitting sources of large ϵ-entropy if arbitrarily long storage is allowed. The result $A < \infty$, if true, would be extremely useful. What is done in *Subsection 5* is to prove $A = B$, and thus reduce the general problem to a graph-theoretic problem.

5. Storage Constants are Equal to Graph-Storage Constants

This subsection shows that the storage constants are equal to their corresponding graph-storage constants. More specifically, we have the following result.

Theorem. The storage constants $A_\lambda^{(n)}$, $\underline{A}_\lambda$, A_λ, $A^{(n)}$, $\underline{A}$, A are equal to their corresponding graph-storage constants $B_\lambda^{(n)}$, $\underline{B}_\lambda$, B_λ, $B^{(n)}$, $\underline{B}$, B for every $\lambda > 0$ and positive integer n.

Proof. We first prove that

$$A_\lambda^{(n)} = B_\lambda^{(n)} \tag{22}$$

for every $\lambda > 0$ and positive integer n. The inequality

$$B_\lambda^{(n)} \leq A_\lambda^{(n)} \tag{23}$$

is immediate, since the definition of $A_\lambda^{(n)}$ given in Eq. (8) takes a supremum over a larger class of probabilistic metric spaces than the definition of $B_\lambda^{(n)}$ given in Eq. (9). The hard result is to prove that

$$A_\lambda^{(n)} \leq B_\lambda^{(n)} \tag{24}$$

and this we now proceed to do.

The condition given as Eq. (24) can be translated as follows: Given n a positive integer, $\lambda > 0$, $\epsilon > 0$, $\rho > 0$, X a probabilistic metric space with $\infty > H_\epsilon(X) > \lambda$, then there exists a graph G such that $H(G) > \lambda$ and

$$\frac{H_\epsilon(X)}{H_\epsilon[X^{(n)}]} \leq \frac{H(G)}{H[G^{(n)}]} + \rho \tag{25}$$

To show this, we claim that it suffices to assume that X is compact. According to Ref. 6, Theorem 1.4, given any $\xi > 0$, there exists a compact subset X_ξ of X such that

$$\mu(X - X_\xi) \leq \xi$$

$$\mu[X^{(n)} - X_\xi^{(n)}] \leq \xi$$

and, by Theorem 4 of Ref. 1, given any probabilistic metric space X and $\epsilon > 0$, $\nu > 0$, such that $H_\epsilon(X)$ is finite, there exists a $\xi > 0$ with the property that if Y is a closed subset of X such that $\mu(X) \leq \xi$, then

$$|H_\epsilon(X) - H_\epsilon(X)| < \nu$$

Since $\infty > H_\epsilon(X) > \lambda > 0$, and, hence, $H_\epsilon[X^{(n)}] > 0$ in our application, we can guarantee by choice of ξ sufficiently small that

$$\left| \frac{H_\epsilon(X_\xi)}{H_\epsilon[X_\xi^{(n)}]} - \frac{H_\epsilon(X)}{H_\epsilon[X^{(n)}]} \right|$$

is less than ρ. Hence, there is no loss of generality in assuming that X itself is compact, and we do so for the remainder of the proof.

First note that any ϵ-partition of X, or of $X^{(n)}$, to be considered can be assumed finite. If $\mathcal{U} = \{U_i\}$ is an ϵ-partition of a compact metric space, such that $\mu(U_i)$ is a non-increasing function of i, then $\mathcal{V} = \{V_i\}$ is an ϵ-partition with the property that

$$\bigcup_{j=1}^{k} V_j$$

is closed for all k, and such that $H(\mathcal{V}) \leq H(\mathcal{U})$, if we define $V_1 = \bar{U}_1$ and, for $i \geq 1$,

$$V_{i+1} = \left(\bigcup_{j=1}^{i+1} \bar{U}_i \right) - \bigcup_{j=1}^{i} V_i$$

Compactness applied to the open covering

$$\left\{ X - \bigcup_{j=1}^{i} V_i \right\}$$

then guarantees that only finitely many V_i are non-empty and, therefore, only finite ϵ-partitions will be considered in the remainder of this article.

It is known (Ref. 1, Theorem 4) that $H_\epsilon(X)$ is continuous from above in ϵ for any X. Let τ be so small that

$$H_\epsilon(X) \leq H_{\epsilon+2\tau}(X) + \eta \tag{26}$$

and let $\mathcal{J} = \{J_1, J_2, \cdots, J_k\}$ be a finite τ-partition of X, which exists because of compactness. Let $g_i \in J_i$ $(1 \leq i \leq k)$, and let K be the finite metric space of the g_i with metric inherited from X. Let a probability measure μ_τ be put on K by

$$\mu_\tau\{g_i\} = \mu(J_i), \qquad 1 \leq i \leq k$$

so that K is a finite probabilistic metric space.

It will be convenient to assume that $\mathcal{U}$ is an ϵ-partition of X with $H(\mathcal{U}) = H_\epsilon(X)$ and such that the boundary of each $U_i \in \mathcal{U}$ has probability 0. If this does not happen, we consider, for small σ, the sets $U_i^{(\sigma)}$ consisting of all points at a distance of at most $\sigma/2$ from some point of U_i. There exist arbitrarily small σ such that the boundaries of all $U_i^{(\sigma)}$ have probability 0, since these boundaries for the same i and different σ are disjoint. An $(\epsilon + \sigma)$-partition can then be culled from among the $U_i^{(\sigma)}$ as before; the continuity of $H_\epsilon(X)$ from above in ϵ then shows that we can assume that the original boundaries all had probability 0.

We now claim that

$$H_\epsilon(K) \leq H_\epsilon(X) + \eta$$

for τ sufficiently small, depending on $\eta > 0$. If $\mathcal{U}_K = \{U_{i;K}\}$ is the partition $\mathcal{U}$ induces on K by putting

$$j_i \in U_{i;K} \Leftrightarrow j_i \in U_i$$

then $\mu_\tau(U_{i;K})$ differ from $\mu(U_i)$ by at most the probability of the set of points at distance at most τ from the boundary of U_i. Since this boundary has probability 0, we conclude that τ can be chosen so small that

$$H(\mathcal{U}_K) \leq H(\mathcal{U}) + \eta$$

and thus

$$H_\epsilon(K) \leq H_\epsilon(X) + \eta \tag{27}$$

Actually, we do not need Ineq. (26) for X, but for $X^{(n)}$ in the form

$$H_\epsilon[K^{(n)}] \leq H_\epsilon[X^{(n)}] + \eta \tag{28}$$

We now need an inequality going the other way. Let $\mathcal{U}_K = \{U_{i;K}\}$ be an ϵ-partition of K achieving $H_\epsilon(K)$, and let $\mathcal{U}$ be the $(\epsilon + 2\tau)$-partition of X that U_K induces by the following process. Define

$$\tilde{U}_i = \bigcup_{j_i \in U_{i;K}} J_i$$

Then, since the J_i have diameters of at most τ, $\tilde{U}_i$ is indeed of diameter of $\epsilon + 2\tau$ at most. Also, the $\tilde{U}_i$ cover

X, and

$$\mu(\tilde{U}_i) \geqq \mu_\tau(U_{i;K}) \tag{29}$$

Hence, if the indexing is such that $\mu_\tau(U_{i;K})$ is non-increasing in i, and if we define $U_1 = \tilde{U}_1$ and, for $i > 1$,

$$U_{i+1} = \tilde{U}_{i+1} - \bigcup_{j=1}^{i} U_j \tag{30}$$

we have

$$\mu\left(\bigcup_{j=1}^{i} U_j\right) \geqq \mu_\tau\left(\bigcup_{j=1}^{i} U_i\right) \tag{31}$$

for all i. Thus, the $(\epsilon + 2\tau)$-partition $\mathcal{U} = \{U_i\}$ of X satisfies (Ref. 1, Lemma 2)

$$H(\mathcal{U}) \leqq H(\mathcal{U}_K) \tag{32}$$

In other words

$$H_{\epsilon+2\tau}(X) \leqq H_\epsilon(K) \tag{33}$$

We then conclude from Ineq. (26) that

$$H_\epsilon(X) \leqq H_\epsilon(K) + \eta \tag{34}$$

which is the desired inequality.

Since $H_\epsilon[X^{(n)}]$ is not 0, because otherwise $H_\epsilon(X)$ would also be zero, Ineqs. (28) and (34) allow us to choose $\eta > 0$ so small that

$$\left| \frac{H_\epsilon(X)}{H_\epsilon[X^{(n)}]} - \frac{H_\epsilon(K)}{H_\epsilon[K^{(n)}]} \right| < \rho \tag{35}$$

for given $\rho > 0$.

Now we construct G. By making an arbitrarily small change in $H_\epsilon(K)$ and $H_\epsilon[K^{(n)}]$, we can demand that the probabilities of all the points in K be rational with the same denominator N; call the probabilistic metric space so obtained L with probability measure μ_L. We have

$$\left| \frac{H_\epsilon(X)}{H_\epsilon[X^{(n)}]} - \frac{H_\epsilon(L)}{H_\epsilon[L^{(n)}]} \right| \leqq \rho \tag{36}$$

Now make L into a graph G with equally likely vertices as follows: If $\mu_L(j_i) = a/N$, a a non-negative integer, replace j_i by a set of a vertices. Connect two distinct vertices by an edge if and only if they arose either from the same j_i, or from different j_i at a distance of at most ϵ. Call the resulting graph G. Then clearly

$$H(G) = H_a(L), \qquad H[G^{(n)}] = H_\epsilon[L^{(n)}]$$

and so G will do for the graph required by Ineq. (25). This proves that $A_\lambda^{(n)} \leqq B_\lambda^{(n)}$, and so

$$A_\lambda^{(n)} = B_\lambda^{(n)} \tag{37}$$

The proof that $A_\lambda = B_\lambda$ cannot be based on Eq. (37), but must be done as follows: In Ineq. (36), choose X, n so that

$$\left| \frac{nH_\epsilon(X)}{H_\epsilon[X^{(n)}]} - A_\lambda \right| < \eta$$

say, with $H_\epsilon(X) > \lambda$. Now

$$\frac{nH_\epsilon(L)}{H_\epsilon[L^{(n)}]} \leqq \frac{H_\epsilon(L)}{\bar{H}_\epsilon(L)} \leqq B_\lambda$$

so that

$$A_\lambda - B_\lambda < \rho + \eta$$

Since ρ, η are arbitrary, we conclude $A_\lambda \leqq B_\lambda$. Recalling that $B_\lambda \leqq A_\lambda$ proves that $A_\lambda = B_\lambda$.

The remaining equalities follow from the ones proved by the definitions. This completes the proof of the theorem.

The storage constant A, as has been stated, remains unknown at the present time, and, for all we know, is anywhere in the extended interval $[1, \infty]$. However, we do have a sequence $\{G_k\}$ of graphs such that

$$H(G_k) - \bar{H}(G_k) \to \infty \text{ as } k \to \infty$$

Determination of the storage constant is left as an open problem.

6. The Deterministic Case

We close this article with some remarks on the deterministic case. For a compact metric space X (or a graph G), define the ϵ-entropy $E_\epsilon(X)$ (or the entropy) as the minimum of the logrithms of the number of sets in coverings of the space by ϵ-sets (Ref. 7, Chap. 1). Product partitions are used as before to define storage constants. Here, the interpretation is that outcomes of X must be transmitted and nothing can be left out. Furthermore, words of constant length must be used for a given partition. Then $E_\epsilon(X)$ is continuous in ϵ from above, and the analogue of the theorem holds for these entropies too.

Similarly, define the ϵ-capacity $C_\epsilon(X)$ (or the capacity $C(G)$ of a graph G) as the logarithm of the maximum number of points that can exist in X at separation greater

than ϵ (Ref. 5, Chap. 4, and Ref. 7, Chap. 1); $C_\epsilon(X)$ is also continuous from above in ϵ. Also

$$C_\epsilon[X^{(m+n)}] \geqq C_\epsilon[X^{(m)}] + C_\epsilon[X^{(n)}]$$

and the absolute ϵ-capacity $\bar{C}_\epsilon(X)$ can be defined. This definition is essentially that of Shannon for the error-free capacity of certain channels (Ref. 5, p. 38, and Ref. 8). Note that

$$C_\epsilon(X) \leqq E_\epsilon(X)$$

and so

$$\bar{C}_\epsilon(X) \leqq \bar{E}_\epsilon(X)$$

The analogue of the theorem can be proved for capacity as well as entropy, and, in fact, for questions about entropy and capacity simultaneously.

We have seen that important questions in metric spaces are really questions about finite graphs. It may well be, however, that this fact does not make the problems any easier to solve.

References

1. Posner, E. C., Rodemich, E. R., and Rumsey, H., Jr., "Epsilon Entropy of Stochastic Processes," *Ann. Math. Statist.*, Vol. 38, pp. 1000–1020, 1967.
2. Fano, R. M., *Transmission of Information*, MIT Press, Cambridge, Mass., 1961.
3. Mowshowitz, A., *Entropy and the Complexity of Graphs*, Ph.D. Dissertation, University of Michigan, 1967.
4. Öre, O., "Theory of Graphs," *Amer. Math. Soc. Colloq. Publ.*, Vol. XXXVIII, Providence, R. I., 1962.
5. Berge, C., *The Theory of Graphs*, Translated by A. Doig, Methuen, London, 1962.
6. Prokhorov, Yu. V., "Convergence of Random Processes and Limit Theorems in Probability Theory," (in translation), *Theory Prob. Appl.*, Vol. I, pp. 157–214, 1956.
7. Vituskin, A. G., *Theory of the Transmission and Processing of Information*, (in translation), Pergamon Press, New York, 1961.
8. Shannon, Claude E., "The Zero-Error Capacity of a Noisy Channel," *IRE Trans. Inform. Theory*, Vol. IT-2, pp. 8–19, 1956.

XX. Communications Elements Research

TELECOMMUNICATIONS DIVISION

A. Spacecraft Antenna Research: Radiation From a Turnstile Antenna Located in a Plasma Shell, R. Woo

1. Introduction

The geometry of interest is shown in Fig. 1. The problem of radiation from a horizontal dipole within a plasma shell was formulated in SPS 37-50, Vol. III, pp. 312–316. The formulation has been extended to a turnstile antenna $\lambda/4$ above an infinite ground plane in the same way as was described in SPS 37-49, Vol. III, pp. 346–355. The radiation patterns have been evaluated with the aid of the IBM 7094 computer, and the results are discussed in this article.

2. Results and Discussion

As was seen in SPS 37-47, Vol. III, pp. 247–257, there are two types of electron density profiles in the wake. In the near wake, electron density increases and then decreases for increasing radial distance. In the far wake, electron density is maximum on axis and decreases for increasing radial distance. These two types of electron density profiles are approximated by the plasma shell configuration of Fig. 1, and the radiation patterns are discussed separately.

a. Far wake. Since electron density decreases radially, the relative permittivity ϵ_1 of region I is greater than ϵ_2 of region II. Radiation patterns for two cases are shown in Fig. 2. When these patterns are compared with those obtained for the plasma column configuration (SPS 37-49, Vol. III, Figs. 37 and 38, p. 351), it is seen that the addition of the plasma shell smoothes the plasma column patterns, as is expected. The respective critical cone angles calculated from Snell's law and applied to the denser medium region I are also shown in Fig. 2. It is seen that the patterns begin their drop in the proximity of these critical cone angles.

b. Near wake. For this configuration the relative permittivity ϵ_1 of region I is less than ϵ_2 of region II. Considerable difficulty was encountered in the evaluation of the radiation patterns on the computer. The difficulties were traced to truncation error and were removed by rewriting the program in double precision and by solving the fields in matrix form on the computer.

The patterns in Fig. 3 are typical of those obtained. The critical cone angle calculated from application of Snell's law to the denser medium region II is also shown. As in the far-wake region, the patterns decrease rapidly with decreasing cone angle for cone angles less than θ_{cr}.

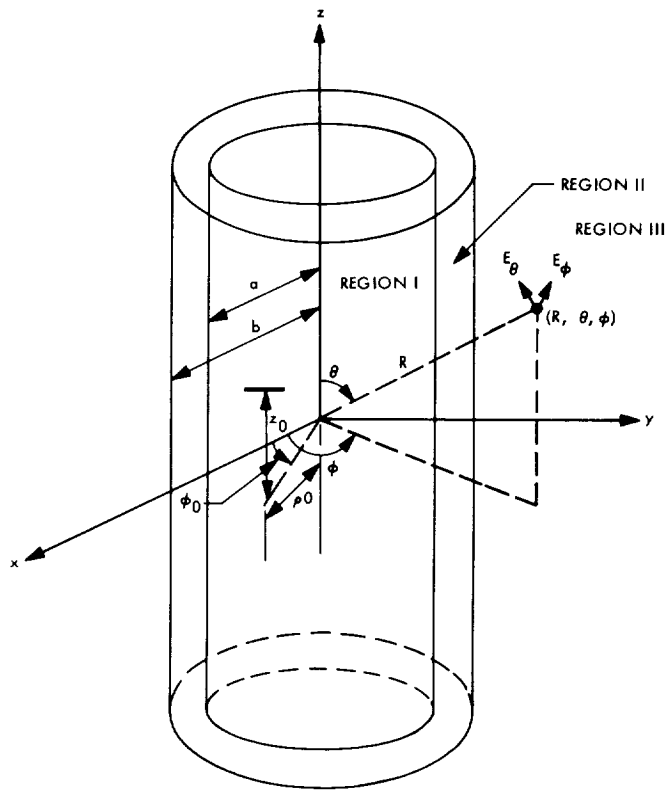


Fig. 1. Geometry of the problem

In addition, there are sharp peaks present. These peaks are due to leaky wave radiation and have been observed in other studies (Refs. 1–3). The physical picture can be described as follows. If radius b is extended to infinity, surface waves would be excited along interface a. When b is finite, the energy that would propagate along interface a as a surface wave if b were infinite, continues to be trapped. However, part of this energy continuously leaks away at interface b as the surface wave propagates. Since the surface wave decays exponentially in the radial direction before it appears at interface b, its amplitude is small, and the leakage of energy is also small. Since the surface wave is a fast wave, the leaky wave pole can be very close to the saddle point, resulting in a sharp peak in the radiation patterns.

Patterns other than those shown in Fig. 3 were obtained. A summary of the leaky-wave beam data from all these patterns are shown in Table 1.

Table 1. Summary of leaky-wave beam data for $\epsilon_1 = 0.9$ and $\epsilon_2 = 0.6$

ka	kb	Location of peak: cone angle, deg	Gain of peak, db
4	6	31.72	12.52
4	8	31.63	18.67
4	10	31.58	24.79
4	12	31.58	30.84
2	10	39.4	16.33
6	10	26.52	23.0
8	10	23.7	16.9
6	8	26.4	15.2
10	12	22.2	17.98

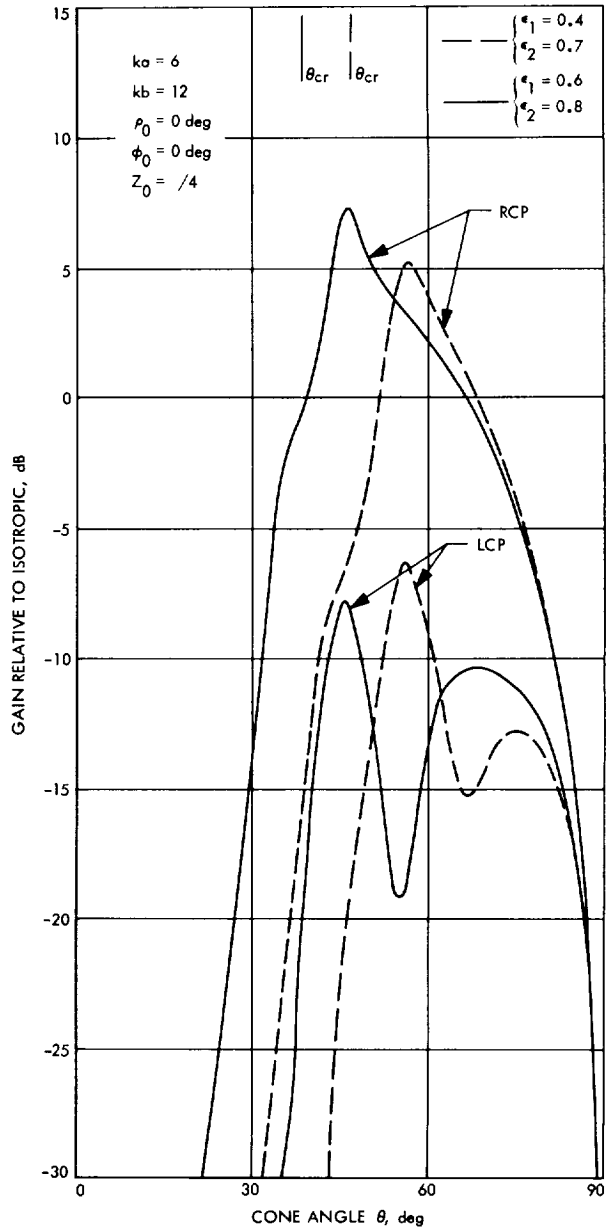


Fig. 2. Radiation patterns for far-wake region

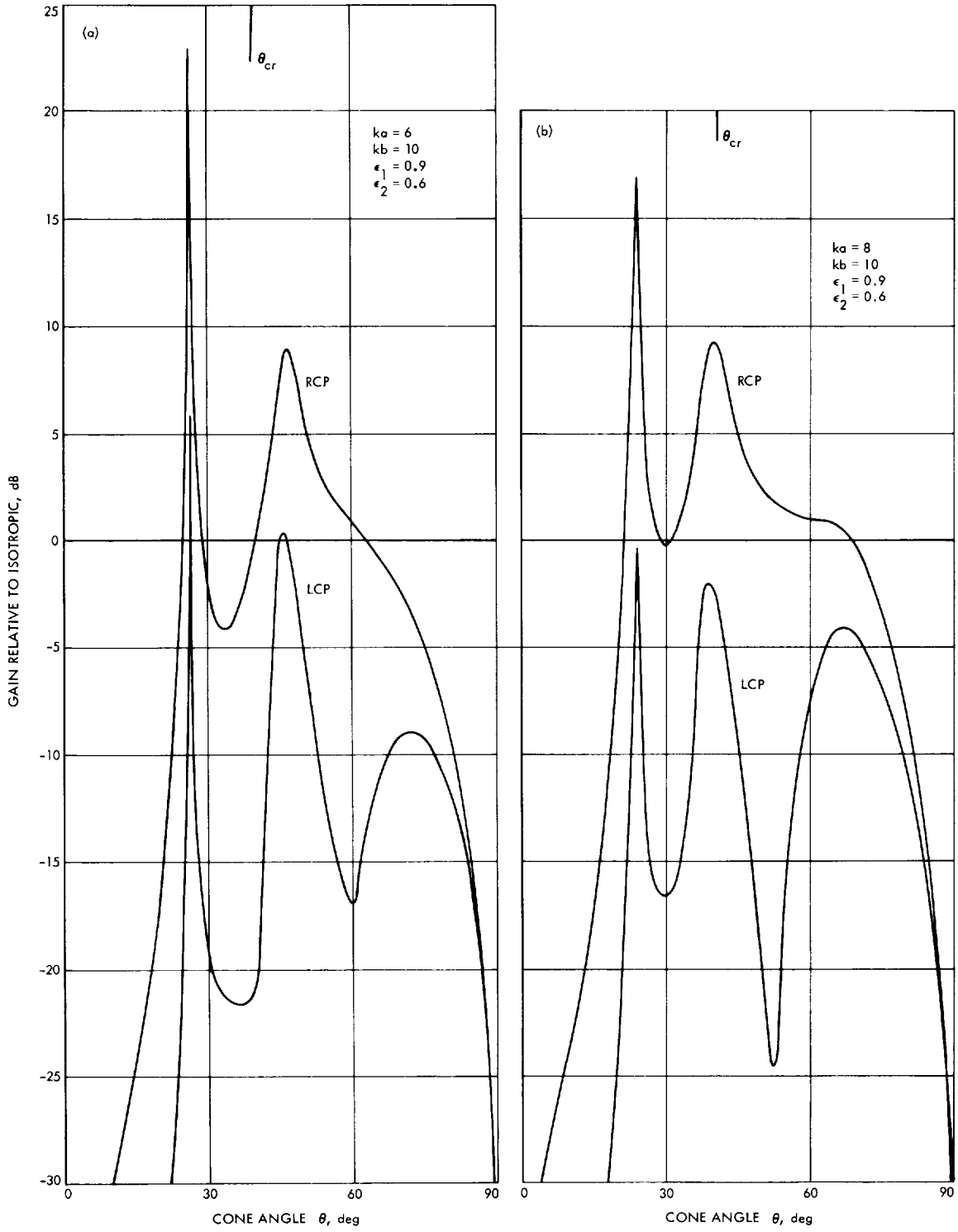


Fig. 3. Radiation patterns for near-wake region: (a) $ka = 6$, $kb = 10$, $\epsilon_1 = 0.9$, $\epsilon_2 = 0.6$; (b) $ka = 8$, $kb = 10$, $\epsilon_1 = 0.9$, $\epsilon_2 = 0.6$

Qualitative results in agreement with those observed in the planar geometry (Ref. 3) are:

(1) If a is fixed and large compared to wavelength, the beam position is approximately independent of b.

(2) When the beam is narrow, the power contained in the beam remains approximately constant. Therefore, the higher the beam rises, the narrower it is.

It should be noted that if the location of a fixed thickness plasma shell is allowed to increase radially, the beam location shifts to small cone angles, and the beam becomes sharper.

The patterns in Figs. 2 and 3 are for the case where the antenna is located on axis. If the antenna is moved off axis, the general effects are similar to those observed in the plasma column case (SPS 37-49, Vol. III, pp. 346–355). However, in the near-wake case, if the antenna is located off axis, higher order surface wave modes will be excited so that additional peaks at cone angles less than θ_{cr} will appear. If ka and kb (k is the free-space wave number) are increased, the effects would be similar to those observed in the plasma column patterns; there will be more ripples in the patterns, and the drop in the proximity of θ_{cr} will be sharper. Additional peaks at cone angles less than θ_{cr} may also appear, since higher order surface wave modes may be excited.

3. Conclusions

The presence of the plasma will cause the patterns to develop an on-axis "null." The extent of the null is proportional to electron density. The null region starts at approximately the critical angle calculated from Snell's law corresponding to the maximum electron density in the wake. The effect of this null is to increase the blackout time for communications cone angles within this null region. In the nonnull region there are no significant depolarization effects, and satisfactory communications may still be carried out.

References

1. Ishimaru, A., "The Effect of Unidirectional Surface Waves Along a Perfectly Conducting Plane on the Radiation From a Plasma Sheath," *Electromagnetic Aspects of Hypersonic Flight*, pp. 147–168. Edited by W. Rotman, H. K. Moore, and R. Papa. Spartan Books, Baltimore, Md., 1964.

2. Harris, J. H., Villeneuve, A. T., and Broca, L. A., "Radiation Patterns From Plasma Enclosed Cylindrical Hypersonic Vehicles," *Radio Sci. J. Res.* NBS 69D, No. 10, pp. 1335–1344, Oct. 1965.

3. Harris, J. H., "Leaky-Wave Beams of Multiply Layered Plasma Media," *Radio Sci.*, Vol. 3, New Series 1, No. 2, pp. 181–189, Feb. 1968.

B. Spacecraft Antenna Research: Antenna Tolerances, *R. M. Dickinson*

1. Introduction

The object of this study is to increase the accuracy of full-scale spacecraft antenna pattern measurements. Antenna patterns recorded from full-scale spacecraft antenna models are used in communications analysis and prediction.

Previous articles have mainly reported instrumentation and measuring techniques. This report will present the results of a preliminary investigation into the feasibility of analytic determination of antenna patterns.

2. Calculated Antenna Patterns

The analytical calculation of antenna patterns will never replace the actual measurement because of the need to verify the theory and accuracy of pattern calculations. However, the capability of calculating patterns should aid in producing more accurate pattern data.

Calculated patterns may be the most economical way to obtain patterns that are representative of the spacecraft antennas in the free-space environment. Present measurement techniques are currently limited, not by instrumentation, but by the ability to simulate free space when the test antenna must be positioned on some support structure near the earth. Reflections and diffraction from surrounding objects and mechanical deflections caused by wind and gravity cause measurement errors.

The two extremes of the problem of obtaining greater accuracy pattern data analytically are exemplified by very high-gain and very low-gain spacecraft antennas.

a. High-gain spacecraft antennas. Because of the limited packaging volume available on spacecraft launch vehicles, the high-gain antennas will most likely be unfurlable types. The lightweight versions of unfurlable antennas could be constructed in such a way that they would not be capable of self support in a 1-g field. However, a tightly furled antenna could survive launch more easily. Therefore, pattern data obtained by measurements on earth presents many mechanical problems. The normal pattern measurement of large-diameter solid antennas is also difficult because of the long distances required.

However, if both the surface equation of the erected surface and the primary illuminator pattern are known, the secondary pattern can be calculated to within only a slight degree of uncertainty by surface current integration techniques (SPS 37-47, Vol. III, pp. 242–247).

The uncertainties involve: (1) possible gain reduction, (2) polarization changes caused by reflector reaction on the feed, and (3) gain loss and sidelobe shape changes due to feed support blockage. In addition, the accuracy of the surface current technique is limited somewhat by the requirement for the radii of curvature of the reflecting surface to be large compared to a wavelength. The erectable antenna reflecting surface is not a smooth figure of revolution.

Thus, techniques are available for calculating the patterns of high-gain spacecraft antennas.

b. Low-gain spacecraft antennas. Techniques for calculating the patterns of low-gain antennas are almost nonexistent, with the exception of the simple half-wave dipole (Ref. 1). Also, when a low-gain antenna is placed on the tortuous geometry of a spacecraft such as *Surveyor*, the problem of calculating the resulting pattern is even more hopeless.

However, the previously mentioned surface current integration techniques can be employed for spacecraft with uncluttered surfaces (such as the later *Mariners*), using medium gain (5 to 8 dB) low-gain antennas. The low-gain antenna is considered to be the primary illuminator, and the spacecraft solar panels are considered to be the reflector.

The major problem is that of obtaining the primary illuminator pattern.

c. Wire antennas. Recent developments (Refs. 2 and 3) in applying integral theory techniques to radiators small in comparison to a wavelength have possible application to the problem of calculating low-gain or primary illuminator antenna patterns. The continuous conducting surfaces of the antenna are modeled by wire segments. The integral equation technique could then be used to determine the currents in the wire segments so that the boundary condition of the tangential electric field equal to zero everywhere except at the feed point is satisfied. The radiation pattern can then be calculated from the current distribution in the wires.

Existing programs (Ref. 2) for calculating scattering (not self-induced radiation) patterns are currently limited to approximately 100 wire segments on an IBM 7094 computer.

An experiment was made to determine the minimum number of wires required to model a typical S-band (2300 MHz) low-gain antenna. Figure 4 shows the solid conductor coaxial cavity antenna and three wire mesh models. The mesh spacings are 0.25, 0.50, and 0.75 in.

Figure 5 shows the measured patterns for the four antennas. The pattern shapes of the solid, 0.25, and 0.5-in. mesh antennas are all approximately the same. The 0.75-in. mesh pattern is very different, exhibiting dipole-like radiation directly from the feed points interfering with the normal pattern shape of the more solid configuration.

Table 2 summarizes the characteristics of the antenna models. The decreasing peak gain with increasing mesh spacing is due to I^2R losses in the wires in the 0.25- and 0.5-in. models, since the pattern shape stayed the same as the solid model. In order to yield the same pattern shape the effective current distribution must remain the same. In the finitely conducting wire models the current is constrained to flow in a much smaller surface area. This increases the current density and, hence, the I^2R losses. The gain losses of the 0.75-in. mesh model are due to both I^2R losses and leakage loss through the widely spaced mesh. Based upon the experimental measurements the maximum wire spacing allowed to model this particular S-band antenna is greater than 0.5 in. but less than 0.75 in. A wire segment spacing of $\lambda/10$ would appear to be a good choice for the maximum allowed spacing. The $\lambda/10$ mesh spacing requirement would result in approximately 500 wires in the computer model of this particular antenna.

However, by modeling simpler feeds than the one illustrated in Fig. 4, and using a computer with greater

Table 2. Antenna model characteristics

Antenna	Number of wire segments	Peak gain, dB	Phase center relative to mouth, in.	Feed point VSWR
Solid conductor	—	7.5	−0.12	2.0
0.25-in. mesh	2200	7.25	+0.08	2.0
0.5-in. mesh	532	5.0	0.0	2.5
0.75-in. mesh	270	2.6	—	5.0

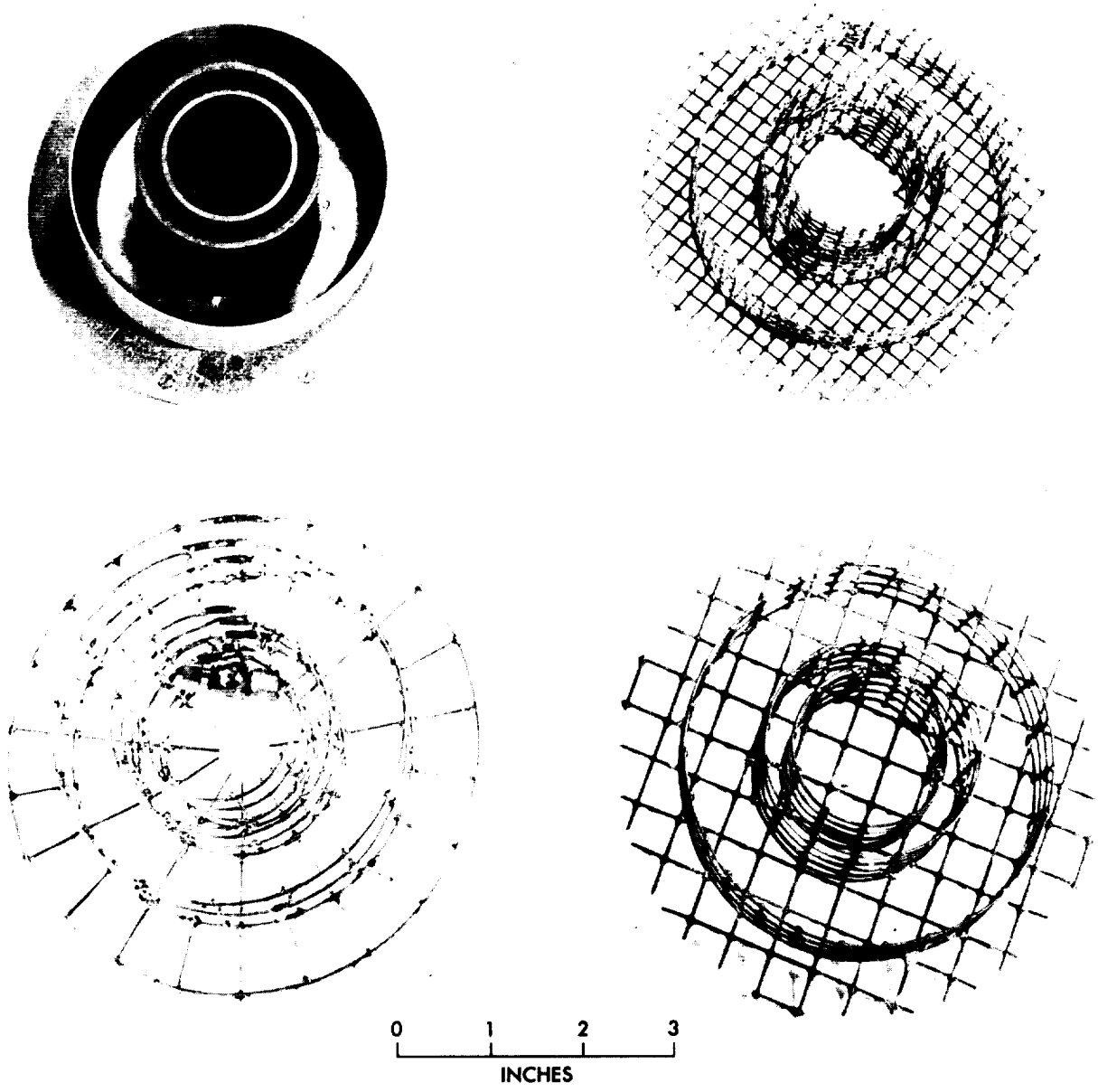


Fig. 4. Coaxial cavity antenna models

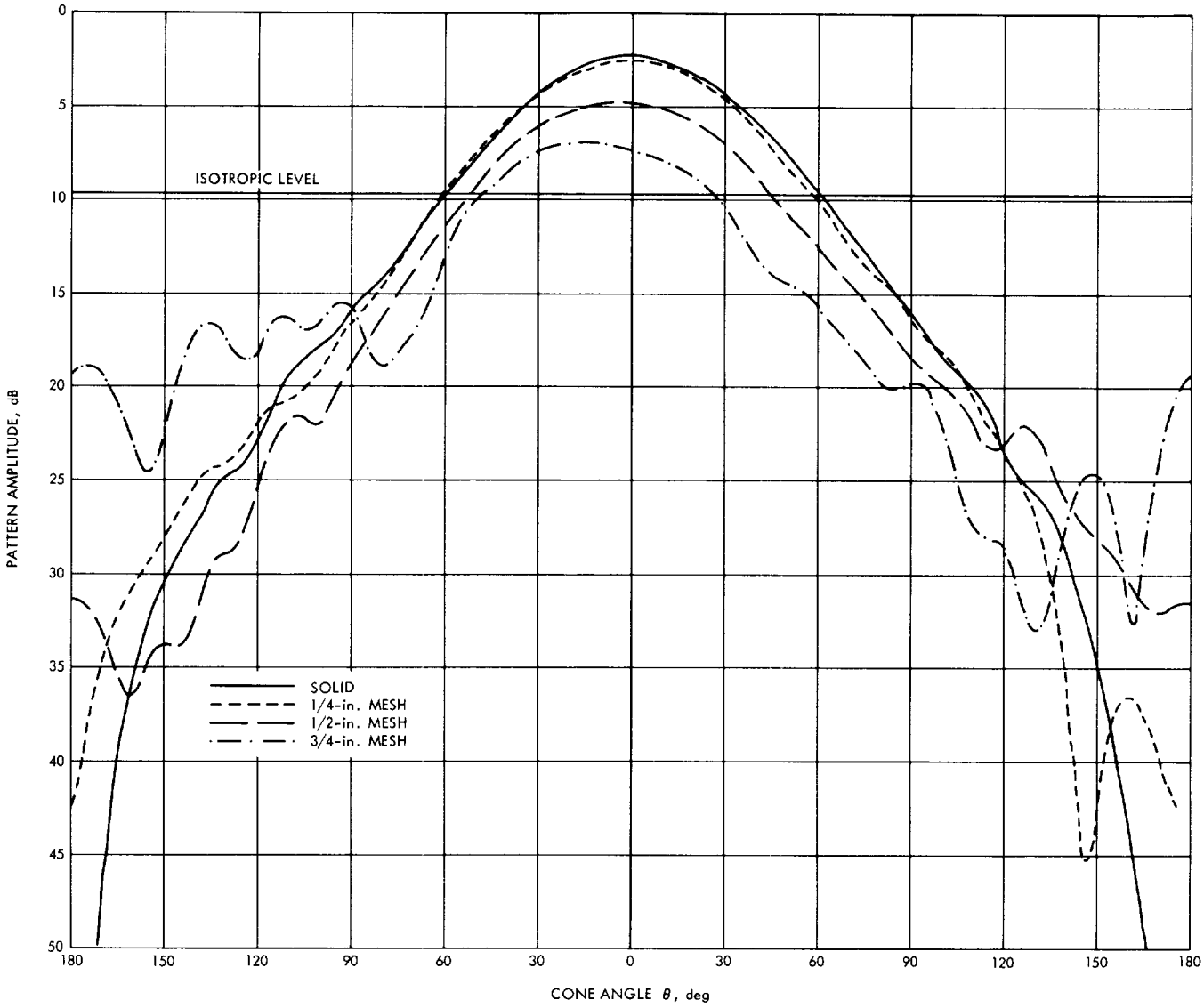


Fig. 5. Coaxial cavity model patterns

speed and core capacity than the IBM 7094, it should be practical to develop a program to calculate the patterns of low-gain spacecraft antennas.

References

1. Kraus, J. D., *Antennas*, Chap. V. McGraw-Hill, New York, 1950.
2. Richmond, J. H., "A Wire-Grid Model for Scattering by Conducting Bodies," *IEEE Transactions on Antennas and Propagation*, Vol. AP-14, No. 6, pp. 782–786, Nov. 1966.
3. Tanner, R. L., and Andreasen, M. G., "Numerical Solution of Electromagnetic Problems," *IEEE Spectrum*, pp. 53–61, Sept. 1967.

C. Precision Calibration Techniques: Microwave Thermal Noise Standards, *C. T. Stelzried*

1. Introduction

Calibrated microwave thermal noise standards (Ref. 1) are used for microwave radiometry (Ref. 2), antenna temperature calibrations (Ref. 3), loss measurements (SPS 37-41, Vol. III, p. 83), low-noise amplifier performance evaluation, and low-level cw signal level calibrations (Ref. 4). A typical thermal noise standard consists of a matched resistive element thermally isolated by a uniform transmission line. The transmission line is usually fabricated from copper-plated stainless steel and has distributed temperatures and transmission loss factors. Solutions of the theoretical noise temperature at the output of a transmission line with various temperature and loss distributions have been tabulated (SPS 37-51, Vol. III, p. 301). A Fortran computer program has been developed for a general solution that uses the transmission-line temperature and loss distributions for input data.

2. Computer Program for Calibration of Microwave Thermal Noise Standards

The available noise power from a termination is given by

$$P = \frac{1}{2} hfB + \frac{hfB}{e^{hf/kT} - 1} \tag{1}$$

where

T = termination temperature, °K

k = Boltzmann's constant, 1.38054×10^{-23} J-°K^{-1}

h = Planck's constant, 6.6256×10^{-34} J-s

B = bandwidth, Hz

f = frequency, Hz

Assuming $hf/kT \ll 1$

$$P = kTB \tag{2}$$

Consider a thermal noise standard, as shown in Fig. 6, consisting of a termination at temperature T and a transmission line with distributed temperatures and propagation constants. Signify the propagating noise power, transmission-line thermal temperature, and propagation constant at x by P_x, T_x, and α_x.

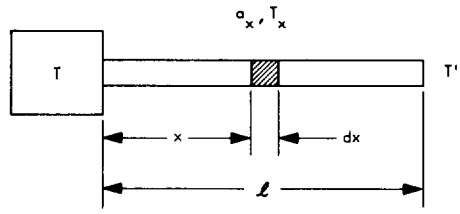


Fig. 6. Thermal noise standard with loss and temperature of transmission line a function of position

The propagating noise power can be separated into two parts. One part is from the termination, attenuated by the transmission line; the second is from the noise contribution of the lossy transmission line. The noise power at the reference output due to the termination is given by P/L, the termination noise power divided by the total line loss. Total line loss is given by

$$L = \exp 2\alpha\ell = \exp\left(\int_0^{\ell} 2\alpha_x dx\right) \tag{3}$$

where ℓ is that portion of the line that has a temperature distribution during operational use. The total noise power at the output reference given by the contribution from the transmission line and the contribution from the termination

$$P' = \frac{2kB}{L}\int_0^{\ell} \alpha_x L_x T_x dx + \frac{P}{L} \tag{4}$$

Dividing by kB gives the noise temperature[1]

$$T' = T'' + \frac{T}{L} \tag{5}$$

[1]IRE Standards on Electron Tubes: Definitions of Terms, 1962 (62 IRE 7.S2), *Proc. IEEE*, March 1963, p. 434.

where

$$T'' = \frac{2}{L}\int_0^{\ell} \alpha_x L_x T_x dx$$

is the contribution from the transmission line.

A Fortran IV computer program (JPL Designation 5847000, CTS34; submitted to Cosmic[2]) has been developed to calibrate microwave thermal noise standards according to Eq. (5). An iteration computing technique is used to transfer the termination temperature T at $x = 0$, $i = 1$ to the output of the transmission line at $x = \ell$, $i = n$ (as shown in Fig. 7). The transmission-line loss distribution is determined by measuring the total loss L_i in decibels at various temperatures $\bar{\bar{T}}_i$. A curve fit is determined from

$$\bar{\bar{L}}_i, \text{dB} = A_1 + A_2 \bar{T}_i + A_3 \bar{T}_i^2 \qquad (6)$$

and the constants $A_1, A_2, \cdots$ are used as input data. The length ℓ' identifies the transmission-line overall length in the above calibration. The operational temperature distribution entered as a table of x_i versus T_i completes the required input data.

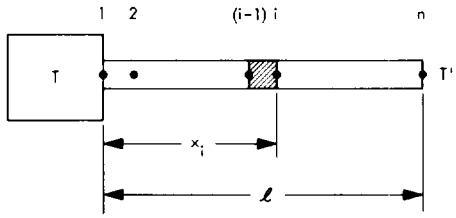


Fig. 7. Thermal noise standard representation for computer program

The thermal temperature of the ith section (from $i - 1$ to i) is estimated by

$$\bar{T}_i = \frac{T_i + T_{i-1}}{2} \qquad (7)$$

The loss of the ith section is then computed from Eq. (6), using the temperature defined by Eq. (7), as follows:

$$\bar{L}_i, \text{dB} = \frac{(\bar{\bar{L}}_i, \text{dB})(x_i - x_{i-1})}{\ell'} \qquad (8)$$

[2]Computer Software Management and Information Center, Computer Center, University of Georgia, Athens.

The loss from the source to output to output of the ith section is calculated by iteration (using an initial value of L_i, dB = 0),

$$L_i, \text{dB} = L_{i-1}, \text{dB} + \bar{L}_i, \text{dB} \qquad (9)$$

It should be noted that $L_n, \text{dB} = L$, dB.

The loss ratio for the ith section is calculated from

$$\bar{L}_i = 1 + \bar{\mathcal{L}}_i + \frac{\bar{\mathcal{L}}_i^2}{2} + \frac{\bar{\mathcal{L}}_i^3}{6} \qquad \bar{L}_i, \text{dB} < 0.02\ \text{dB}$$

$$= {}_{10}\bar{L}_i, \text{dB}/10 \qquad \text{otherwise} \qquad (10)$$

where

$$\bar{\mathcal{L}} = \frac{\bar{L}_i, \text{dB}}{10 \log_{10} e}$$

The noise-temperature contribution of the ith section of the transmission line (assuming linear temperature and linear propagation loss distributions across the section, per case 5, Table 1, SPS 37-51, Vol. III, p. 301) is calculated from

$$\bar{T}_i'' = \bar{T}_i \bar{\mathcal{L}}_i - \frac{1}{6}(T_i + 2T_{i-1})\bar{\mathcal{L}}_i^2 \qquad \bar{L}_i, \text{dB} < 0.02\ \text{dB}$$

$$= \left[1 - \frac{1 - \frac{1}{\bar{L}_i}}{\bar{\mathcal{L}}_i}\right] T_i - \left[\frac{1}{L_i} - \frac{1 - \frac{1}{\bar{L}_i}}{\bar{\mathcal{L}}_i}\right] T_{i-1}$$

otherwise (11)

The total noise contribution due to the transmission-line length from the source to the output of the ith section is calculated by iteration (using an initial value of $T_i'' = 0$),

$$T_i'' = \bar{T}_i'' + \frac{T_{i-1}''}{\bar{L}_i} \qquad (12)$$

It should be noted that $T_n'' = T''$. The total transmission-line loss is obtained from Eq. (9) using

$$L_i = 1 + \mathcal{L}_i + \frac{\mathcal{L}_i^2}{2} + \frac{\mathcal{L}_i^3}{6} \qquad L_i, \text{dB} < 0.02\ \text{dB}$$

$$= {}_{10}L_i, \text{dB}/10 \qquad \text{otherwise} \qquad (13)$$

where

$$\mathcal{L}_i = \frac{L_i,\text{dB}}{10 \log_{10} e}$$

The total noise contribution due to the transmission line and the source defined at i is

$$T'_i = T''_i + \frac{T}{L_i} \tag{14}$$

It should be noted that $T'_n = T'$.

Identification of the computer symbols with the notation is given in Table 3. Various test cases have been generated comparing theoretical solutions (Table 1, SPS 37-51, Vol. III, p. 301) to computer-iterated solutions to provide necessary verification of the program. The number of increments (or increment size) is an important consideration in the computing accuracy. The use of case 5 rather than case 1 in Eq. (11) minimizes the computing error when using large transmission-line increments. A least-squares fit was performed on the input temperature distribution data given in the text. The computing error from using the 11 data points is less than 0.001°K for this example. More data points are not required, considering the other sources of error.

3. Operational Microwave Thermal Noise Standards

The Deep Space Network communications system employs 85- and 210-ft-diam antennas using cassegrain feed configurations. The low-noise transmission-line components, maser amplifier, and calibration equipment are installed in removable (plug-in) feed cones. The standard DSN and experimental research cones are interchangeable among the various antennas. The DSN standard cassegrain cones presently use a WR 430 waveguide ambient and a 0.875-in. coaxial liquid-nitrogen-cooled termination[3] installed in a 20-liter Linde dewar for low-noise calibrations. The accuracy of these calibrations is directly dependent on the accuracy of the noise standard calibrations. Improved noise-temperature calibration accuracy has been achieved in the research cones[4], using all waveguide ambient and liquid-nitrogen-cooled thermal noise standards. The waveguide configuration has lower transmission-line loss, better connector-insertion loss repeatability, and, therefore, can be more accurately calibrated and maintained.

[3]Model SP 9025A, Maury Microwave Corp., Montclair, Calif.

[4]Levy, G. S., *et al.*, "The Ultra Cone: An Ultra Low-Noise Space Communications Ground Receiving System," to be published in *IEEE Trans. Microwave Theor. Tech.*, September 1968.

Table 3. Nomenclature for Fortran IV computer program

Computer printout	Computer program	Text notation	Definition
X(I)	X(I)	x_i	Distance from termination to *i*th data point of line
L'	DLP	ℓ'	Line length for $\bar{\bar{L}}_i$ evaluation
	NCF		Computer input for the number of coefficients in Eq. (6)
LBB, DB(A)	A(NC)	$A_1, A_2, A_3, \ldots$	Constants defining $\bar{\bar{L}}_i$ in Eq. (6)
	NC		Coefficient number in Eq. (6)
LBB(I), DB	LBB(I)	$\bar{\bar{L}}_i$, dB	Loss when at constant temperature $\bar{T}_i$, dB
	LB(I)	$\bar{L}_i$	Loss of line increment (from $i-1$ to i), ratio
LB(I), DB	LBDB(I)	$\bar{L}_i$, dB	Loss of line increment (from $i-1$ to i), dB
	L(I)	L_i	Loss of line from 0 to i, ratio
L(I), DB	LDB(I)	L_i, dB	Loss of line from 0 to i, dB
L, DB	LDB(N)	L	Total line loss, dB
T(I), K	T(I)	T_i	Physical line temperature at *i*th data point, °K
TB(I), K	TB(I)	$\bar{T}_i$	Average physical temperature of line increment (between $i-1$ and i), °K
TB"(I), K	TBPP(I)	$\bar{T}''_i$	Thermal noise contribution due to line increment (between $i-1$ and i) defined at i, °K
T"(I), K	TPP(I)	T''_i	Thermal noise contribution due to line between 0 and i, defined at i, °K
T", K	TPP(N)	T''	Total thermal contribution of line, °K
T'(I), K	TP(I)	T'_i	Calibrated temperature defined at i, °K
T', K	TP(N)	T'	Calibrated temperature, °K
T, K	TT	T	Termination temperature, °K
	SL(I)	$\mathcal{L}_i$	L_i, dB/10 $\log_{10}e$
	SLB(I)	$\bar{\mathcal{L}}_i$	L_i, dB/10 $\log_{10}e$

4. Noise-Temperature Calibration

A commercial termination[5] is used in the standard cassegrain ultra cone (see Footnote 4) in a liquid nitrogen 40-liter dewar. This termination is designed primarily for use at 2295 MHz. The VSWR at 2295 MHz is approximately 1.013. The VSWR bandwidth is limited primarily by the half-wavelength plastic-foam-sealed window in the 30-deg waveguide bend. Techniques are available to broaden the bandwidth, if required (Ref. 5).

[5]Model SR 8135 SN 002, Maury Microwave Corp., Montclair, Calif.

The temperature distribution (Figs. 8 and 9) along the stainless steel thin-wall waveguide section is measured with chromel constantan thermocouples referenced to liquid nitrogen. The data are taken with the dewar in a vertical orientation. A linear temperature distribution along the thin-wall stainless steel waveguide section is assumed. The temperature changes from approximately 2 to 19°K above the reference temperature during the operating liquid levels of the dewar. The total insertion loss of the copper waveguide section was measured to be less than 0.01 dB. This represents less than 0.02°K output noise temperature variation (Case 3, Table 1, SPS 37-51, Vol. III, p. 301) with liquid level.

Temperature measurements were made to verify that the termination material achieves temperature equilibrium with the liquid nitrogen. The thermocouples were imbedded directly in the termination material and referenced to the liquid nitrogen. The temperature difference was not detectable (the instrumentation resolution was better than 0.02°K) even with an infrared heat lamp pointed at the open 15-in. length of copper waveguide. However, it is imperative that the termination section be completely submerged in the liquid nitrogen.

A dissipative loss of (0.0085 $\pm$0.003 pe_{total})dB was measured for the 4.2-in.-length aluminum 30-deg waveguide bend at ambient (20°C) temperature, using the transmission technique (Ref. 6) at 2295 MHz. The transmission-line loss of the stainless steel section was measured in the neighborhood of three separate temperatures, using both transmission (Ref. 7) and cavity resonance (Ref. 8) techniques (Refs. 9, 10, and SPS 37-43, Vol. III, p. 54) with results as indicated in Table 4 and Fig. 10. The measurement temperatures were determined by the ambient temperature (300°K), dry ice and

Table 4. 2295-MHz dissipative loss of stainless steel WR 430 waveguide section versus temperature

Physical temperature, °K	Transmission-line loss at 2295 MHz, dB
299.0	0.0149[a, b]
210.8	0.0052[a]
196.8	0.0045[b]
107.0	0.0016[a]
77.4	0.0017[b]

[a]JPL transmission technique.

[b]Rantec cavity Q technique (translated to 2295 MHz as defined at 299.0°K).

MICROWAVE NOISE STANDARD CALIBRATION: CTS/34 CASE 1DA

	K(I)	T(I),K	LBB(I),DB	LB(I),DB	L(I),DB	TB(I),K	TB''(I),K	T''(I),K	T'(I),K
1	0.000	83.300							
2	0.479	114.500	0.14477E-02	0.11557E-03	0.11557E-03	98.9000	0.0026	0.0026	77.3606
3	0.962	142.300	0.14348E-02	0.11550E-03	0.23108E-03	128.4000	0.0034	0.0060	77.3619
4	1.454	166.500	0.20008E-02	0.16406E-03	0.39514E-03	154.4000	0.0058	0.0119	77.3648
5	1.921	191.100	0.30239E-02	0.23536E-03	0.63050E-03	178.8000	0.0097	0.0216	77.3703
6	2.384	212.900	0.44385E-02	0.34251E-03	0.97300E-03	202.0000	0.0159	0.0375	77.3802
7	2.850	234.600	0.61558E-02	0.47810E-03	0.14511E-02	223.7500	0.0246	0.0621	77.3963
8	3.324	254.600	0.81575E-02	0.64444E-03	0.20955E-02	244.5000	0.0353	0.0984	77.4211
9	3.787	272.100	0.10255E-01	0.79131E-03	0.28869E-02	263.3500	0.0480	0.1464	77.4550
10	4.270	288.900	0.12419E-01	0.99973E-03	0.38866E-02	280.5000	0.0646	0.2109	77.5017
11	4.765	306.300	0.14811E-01	0.12219E-02	0.51085E-02	297.6000	0.0837	0.2946	77.5636
12	5.074	306.300	0.16118E-01	0.83010E-03	0.59386E-02	306.3000	0.0585	0.3530	77.6073
13	5.383	306.300	0.16118E-01	0.83010E-03	0.67687E-02	306.3000	0.0585	0.4115	77.6510

L' = 6.000
T,K= 77.360

L,DB = 0.0059
T'',K= 0.4115
T',K = 77.6510

LBB,DB
(A)
0.65712E-02
-0.91372E-04
0.40007E-06

Fig. 8. Computer printout of thermal noise standard calibration

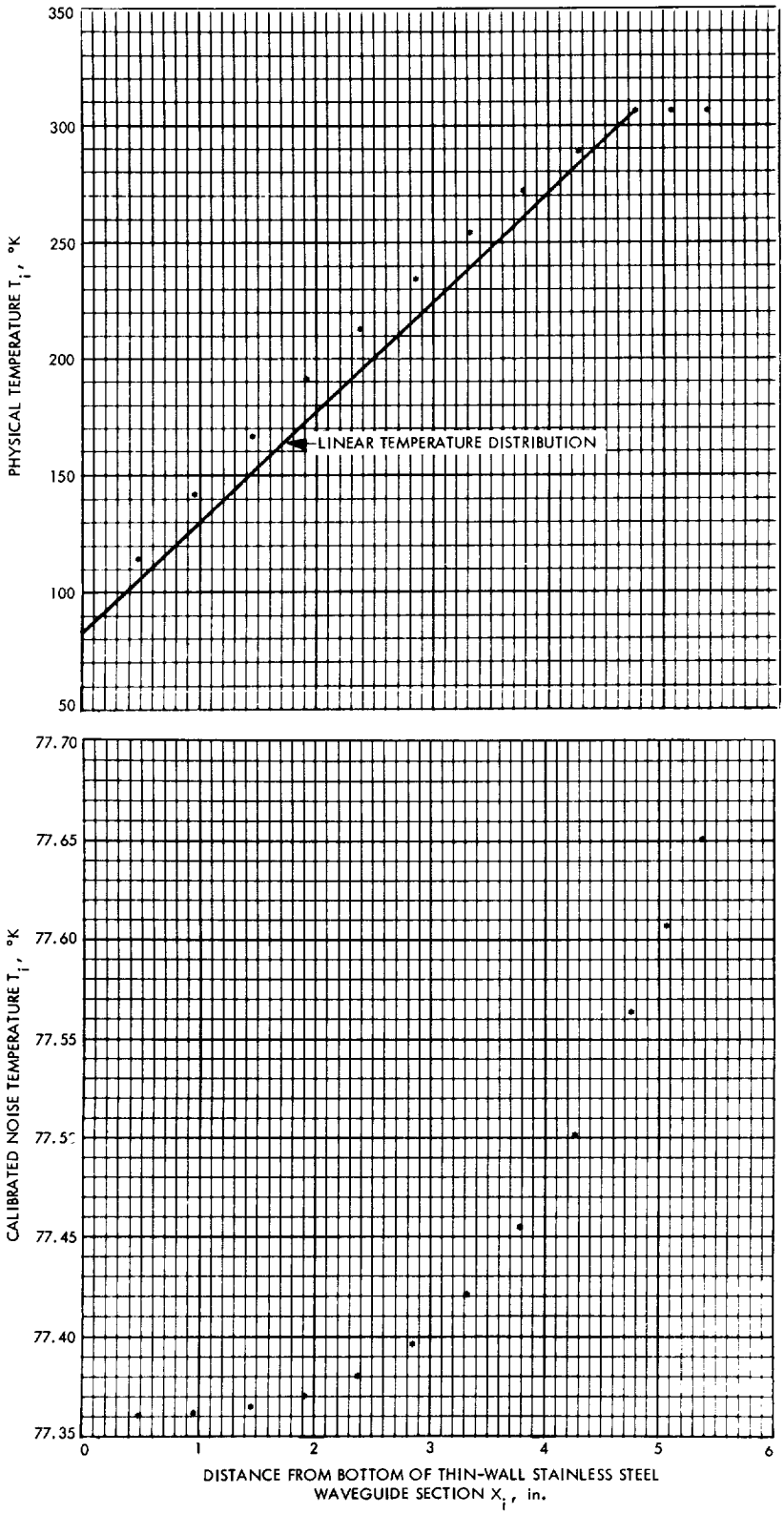


Fig. 9. Physical and calibrated noise-temperature distributions for thermal noise standard

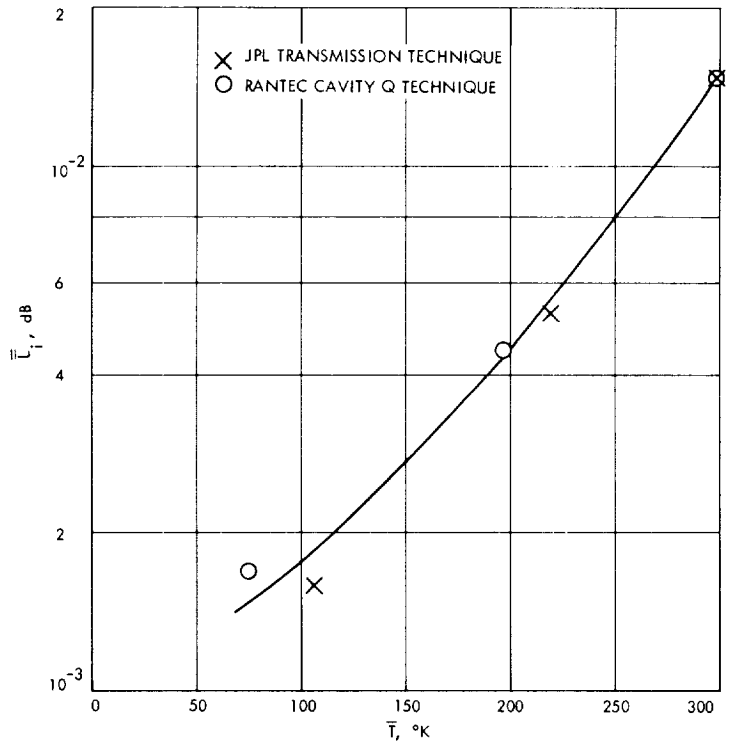


Fig. 10. Insertion loss of stainless steel WR-430 waveguide section versus temperature

alcohol cooling (200°K), and liquid nitrogen cooling (80°K). A second-order curve fit of these values results in the coefficients (Eq. 8):

$$A_1 = 6.5712 \times 10^{-3}$$

$$A_2 = -9.1372 \times 10^{-5}$$

$$A_3 = 4.0007 \times 10^{-7}$$

The loss and temperature distribution measurements are used with the computer program as shown in the printout of Fig. 8 to define the nominal noise temperature at the output of the stainless steel section. Figure 9 shows the measured physical and calibrated noise temperature distributions for the stainless steel waveguide section. The transmission line contributes the major portion of the increased noise temperature at the upper hot end. The noise temperature at the output of the 30-deg bend is computed, assuming a linear temperature distribution from 306.3 to 293.2°K (20°C) and an input temperature of 77.651°K (Fig. 8). The nominal calibrated noise temperature at the output reference flange is 78.09°K. Table 5 indicates calibration errors that result from various measurement errors. These errors are obtained from the computer program with various input temperature distributions, etc. The largest source of error for this particular standard would be obtained if account were not taken for the change in transmission-line loss with temperature. Table 6 summarizes the actual calibration error estimates for this liquid-nitrogen-cooled noise standard. The error estimate for loss measurements is estimated to

Table 5. Noise-temperature errors due to various potential calibration errors at output of stainless steel waveguide section

Parameters	Noise temperature errors, °K
Straight-line temperature distribution (83.3 to 306.3°K)	0.031
10% linearity error in temperature distribution	0.003
5% linearity error in loss measurement	0.013
Constant loss (0.0149 dB at all temperatures)	0.122
1°C error in stainless steel waveguide upper flange temperature	0.003
6°C error in stainless steel waveguide lower flange temperature	0.004

Table 6. Calibration errors for the waveguide liquid-nitrogen-cooled termination[a]

Parameter error	Noise-temperature calibration peak error, °K
Transmission-line loss measurements (assuming $<5\%$ linearity error)	0.034
Transmission-line temperature distribution (assuming 10% nonlinearity)	0.003
Cryogenic liquid pressure (assuming 0.2 mm/Hg error in pressure)	0.020
Termination/liquid temperature difference (resolution of laboratory thermocouple measurement)	0.020
VSWR mismatch (assuming a system with comparable matches)	0.013
Liquid nitrogen level changes (during operational use of approximately 15h)	0.033
Computational error	0.001

[a]Sum = 0.12°K.
rss = 0.06°K.

be between 0.06°K (rss) and 0.12°K (sum). The lower value is probably valid (since most of these errors are not correlated), although the higher value can be considered conservative, to allow for further unresolved sources of error.

5. Conclusion

Precise calibration of operational microwave thermal noise standards requires the use of a computer program. This is because the assumptions made in the theoretical solutions, such as linear temperature and loss distributions, are not exactly satisfied. The calibration error that results from assuming a theoretical linear temperature distribution and an ambient temperature transmission loss results in approximately 0.1°K error for the standard evaluated. This error would be considerably higher with larger transmission-line losses such as are normally encountered when a coaxial transmission line is employed.

The computer program minimizes the effect of large segment size by using either the higher order terms in the loss-calculation expansion or an exact expression for losses greater than 0.02 dB.

With additional effort and expense, increased accuracy in the calibration accuracy of thermal noise standards could be obtained. These improvements include submerging the termination directly in the liquid, using liquid helium (which has a lower sensitivity to pressure), increasing the accuracy in the transmission-line loss and temperature distribution measurements, etc., as indicated by Table 6. Further study could account for such small effects as nonhomogeneous transmission-line loss effects. More effort should be directed toward evaluating and improving the long-term calibration stability of thermal noise standards, especially under field conditions.

References

1. Stelzried, C. T., "A Liquid Helium Cooled Coaxial Termination," *Proc. IRE*, Vol. 49, No. 7, p. 1224, July 1961.
2. Roll, P. G., and Wilkinson, D. T., "Measurement of Cosmic Background Radiation at 3.2-cm Wavelength," *Ann. Phys.*, Vol. 44, p. 289, September 1967.
3. Schuster, D. L., Stelzried, C. T., and Levy, G. S., "The Determination of Noise Temperatures of Large Paraboloidal Antennas," *IRE Trans. Ant. Prop.*, Vol. AP-10, No. 3, p. 286, May 1962.
4. Stelzried, C. T., and Reid, M. S., "Precision Power Measurements of Spacecraft CW Signal Level With Microwave Noise Standards," *IEEE Trans. Instrum. Meas.*, Vol. IM-15, No. 4, p. 318, December 1966.
5. Stelzried, C. T., et al., *Broadband Microwave Waveguide Window*, IR 30-891, Jet Propulsion Laboratory, Pasadena, Calif., April 1966.
6. Stelzried, C. T., Reid, M. S., and Petty, S. M., "A Precision DC Potentiometer Microwave Insertion Loss Test Set," *IEEE Trans. Instrum. Meas.*, Vol. IM-15, No. 3, p. 98, Sept. 1966.
7. Stelzried, C. T., and Mullen, D. L., *Cooled Transmission Line Loss Evaluation Method*, IR 30-1490, Jet Propulsion Laboratory, Pasadena, Calif., June 6, 1968.
8. Document 80083, Rantec Corp., Calabasas, Calif., Feb. 12, 1968.
9. Ginzton, E. L., *Microwave Measurements*, p. 469. McGraw-Hill Book Co., Inc., New York, 1957.
10. Kelly, K. C., *Refrigerated Transmission Line Study, Vol. I, Microwave Analysis and Measurements*, Rantec 66327-FR, Rantec Corp., Calabasas, Calif., Final Report prepared for Jet Propulsion Laboratory under Contract 951638, April 1967. (Scientific and Technical Information Division, NASA Hq., Washington, D.C. 20546, Accession No. N-68-11792.)

D. A Precision DC Potentiometer Insertion Loss Test Set and Reflectometer for Use at 90 GHz,

D. A. Oltmans and T. Sato

1. Introduction

Precision waveguide insertion loss calibrations are critical in many phases of radio-astronomy and communication systems. A dual-channel 90-GHz insertion loss

test set (Ref. 1) has been constructed for these measurements with an absolute accuracy of about 10^{-4} dB at all microwave frequencies.

2. Test Equipment

The insertion loss test set employs two Hewlett Packard Model 431B power meters with power-meter heads designed for use to 40 GHz. Each power meter head requires two thermistors, one for RF detection and the other for thermal compensation.

The insertion loss test set uses a new FXR Model E208A01 dual thermistor mount and operates at 90 GHz (Figs. 11 and 12). A reflectometer was also built on the same aluminum plate to simplify the measurements of the waveguide components' VSWR and the tuning of the heads of the insertion loss set. A special denting tool was used to tune the matching sections.

The RF power for the insertion loss set is generated by a Varian Model VC-713 reflex klystron. A power output of 200 mw (min.) is required for correct operation of the system.

3. Operation

a. Reflectometer. With the switch (Fig. 12) in the most counterclockwise position the RF power is directed into the reflectometer system. The klystron is reflector-modulated with a 1000-Hz square wave to produce the required signal to the standing wave ratio meter. The reflectometer system is of standard configuration using two FXR model (E164A) rotary vane attenuators as an attenuation standard to reduce errors of the individual attenuator at large angles of rotation (i.e., near 90 deg).

Because of the high frequency (i.e., short wavelengths, around 3.3 mm), the screw tuning method used at lower frequencies becomes too unstable at 90 GHz. To greatly reduce this instability a waveguide denting tool (Fig. 13) was used to permanently dent the waveguide in the E-plane only. This achieved the same effect as lowering a screw into the waveguide in lower frequency screw tuners. The denting tool has variable denting rate of from 0.05 to 0.01 in. per knob revolution to make possible coarse and fine tuning. Using a trial-and-error denting technique, reflectometer return losses of 42 dB were easily obtainable by comparing a short to a sliding load.

Fig. 11. Insertion loss test set and reflectometer

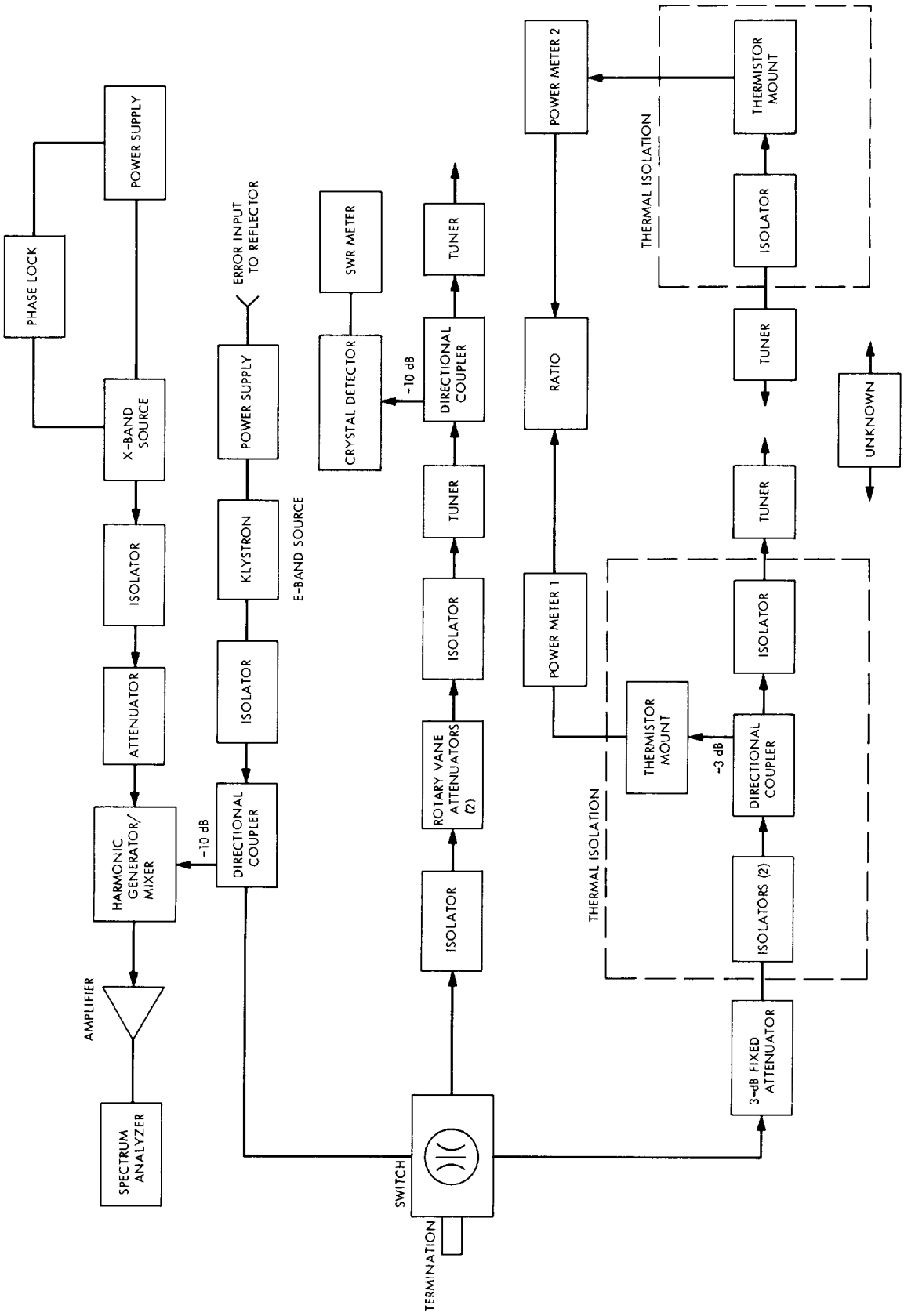


Fig. 12. Insertion loss test set and reflectometer block diagram

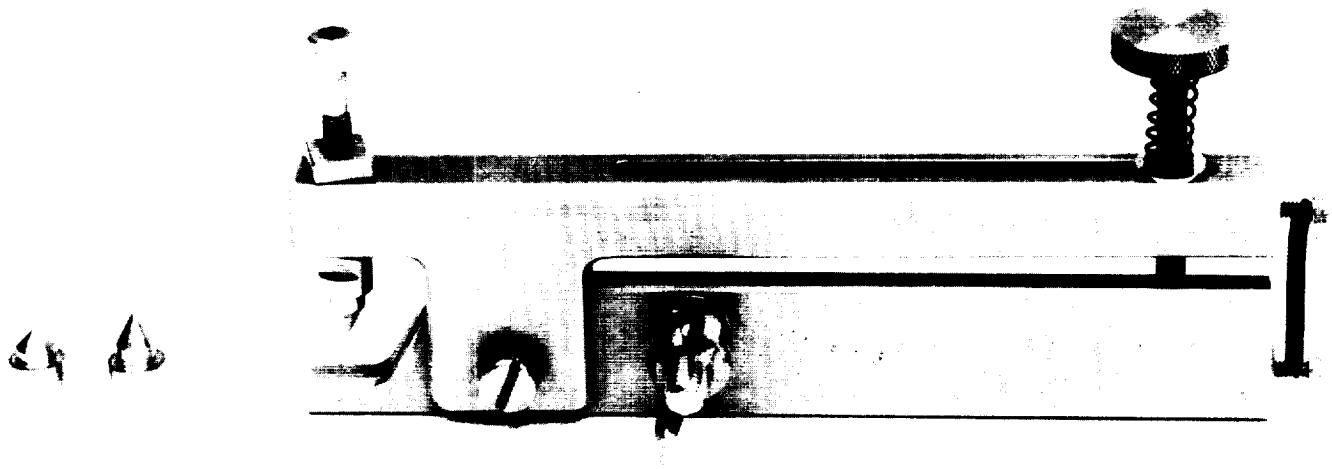


Fig. 13. Denting tool for fabricating E-band waveguide matching sections

Return losses of 46 dB (VSWR = 1.01) or better were achieved, but required many dents. The tuning is reasonably stable from day to day, because the tuning is permanent; however, it is still necessary to fine tune the reflectometer before using it each time for the best performance.

b. Insertion loss test. For insertion loss measurements the klystron was run cw through two Baytron isolators and then into a 3-dB directional coupler. At the output of the auxiliary arm the compensated thermistor mount for power meter 1 (Ref. 2) was connected. At the output of the main arm another isolator and a tuning section were connected. Power meter head 2 (Fig. 14) was preceded first by an isolator and then a tuning section. Both thermistor mounts were thermally isolated in polyurethane foam. The VSWR looking into the open end of each tuning section was optimized by comparing each section to a short at the output of the reflectometer and denting the tuning section until the maximum return loss was obtained. The mounts were adjusted for maximum cw power reading at the desired frequency by positioning the power-detecting thermistor card in its slit and by positioning the sliding short. This adjustment was found to critically affect the efficiency of the mount (Ref. 1). After this was done, the matching sections were tuned, using the previously tuned reflectometer and the denting tool. Return losses were easily tuned to 46 dB or better (46 dB $\approx$ VSWR of 1.01). It was observed that the adjustments of the power thermistor card in the slit and the adjustment of the short affected the match. Therefore, it was important to follow the above order of tuning and to be sure that the cards and shorts were tightly locked in place before tuning the matching sections.

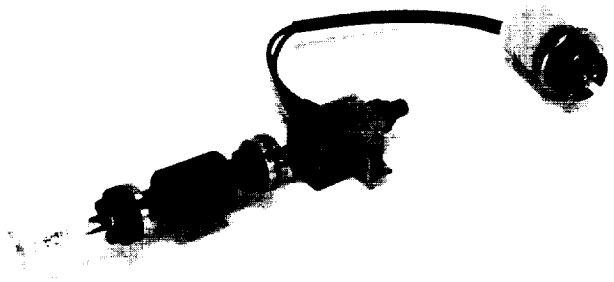


Fig. 14. E-band thermistor mount isolator and VSWR matching assembly

The insertion loss measurements were made by first measuring a reference voltage ratio between power meters 1 and 2 without the test piece (a 3-dB fixed attenuator) between the two tuners and then measuring a new voltage ratio with the test piece in place. The difference between these two ratios in dB was the insertion loss of the test piece, assuming all matches to be perfect. Two sets of nine measurements were made.

4. Performance

The reflectometer was tuned to a 0.3-dB swing for a sliding short and a 42.8 to 46.3 dB return loss for a sliding

load. By comparing the reflection coefficients to those of a second load, the VSWR of the reflectometer was found to be 1.012.

The input and output heads of the insertion loss test set were tuned to a return loss of 46.9 and 49.0 dB, respectively. The match looking into both ends of the test piece was also measured so that reflection losses could be separated from dissipative losses. The insertion loss data and that of the reflectometer measurements were entered into the computer (Ref. 1). Only periodic resetting of signal frequency was made during the first set, whereas the frequency was reset just before each measurement during the second set. The results showed that the standard deviation and probable error were about 50% less when the frequency was reset each time. This was due to the narrow band characteristics of the system caused by the dent tuning.

The attenuator measured insertion loss was calculated to be 3.2086 dB. The average weighted dispersion probable error was calculated to be 0.00178 dB. Added to this were the reflective probable errors due to the reflectometer and mismatch (i.e., system resettability, linearity, and mismatch). This gave a total probable error of 0.00485 dB, which showed that the contribution to probable error due to the reflectometer and mismatch was greater than that of dispersion. Linearity and resettability probable errors could be reduced by calibrating the reflectometer's attenuation standards.

After the insertion loss measurements were made, a recorder was used to show the system's drift and instability. The charts showed that instability was introduced by the signal generator due to frequency drift.

5. Conclusion

It has been demonstrated that with the use of the new FXR model E208A01 dual thermistor mounts, the existing precision DC potentiometer microwave insertion loss test set can be used to make precision insertion-loss measurements at 90 GHz. It is also clear that the present total system probable error of 0.00485 dB can be further reduced to obtain more precise measurements of the waveguide components, especially with lower loss components.

References

1. Stelzried, C. T., et al., "A Precision DC Potentiometer Microwave Insertion Loss Test Set," *IEEE Trans. Instrum. Meas.*, Vol. IM-15, No. 3, p. 98, September 1966.

2. *MM-Wave Power Meter Mount,* IR 30-1141, Jet Propulsion Laboratory, Pasadena, Calif., 1967.

E. Accuracy of Numerically Computed Electromagnetic Scattered Patterns,

S. A. Brunstein, R. E. Cormack, and A. C. Ludwig

1. Introduction

Computer programs have been used extensively to evaluate the performance of existing or proposed Deep Space Net antennas (Ref. 1). A critical link in the evaluation is the computation of the scattering from the subreflector of the cassegrainian configuration used in these antennas. In order to minimize the amount of computer time used by this type of program, data is presented which may be used to ensure accurate results without excessive use of machine time.

The programs covered by this study are: (1) the Rusch scattering program for reflectors which are a figure of revolution (Ref. 2), and (2) the Ludwig program for asymmetrical reflectors.[6] In the Rusch program the scattering surface integral is reduced analytically to a one-dimensional integral, which is evaluated numerically using Simpson's rule. In the Ludwig program the two-dimensional integral is evaluated numerically using the technique described in the above reference.

2. Point Accuracy of Computed Patterns

A typical subreflector feed system is shown in Fig. 15. The same reflector, a figure of revolution azimuthally, was used for testing both programs. The "standard" pattern for this configuration is a pattern computed by the Rusch program. The reasons for accepting this as a valid standard will be discussed in some detail later.

The variable under study is the density of data points used in numerically evaluating the integral. This density will be expressed in terms of the number of data points on a line segment one wavelength long on the surface of the subreflector for a one-dimensional density, or on an area segment one wavelength square for a two-dimensional density. Thus the step size in the integration parameter is translated into a physical distance on the subreflector surface.

[6]Ludwig, A. C., "Computation of Radiation Patterns Involving Double Numerical Integration," to be published in *IEEE Trans. Ant. Prop.*

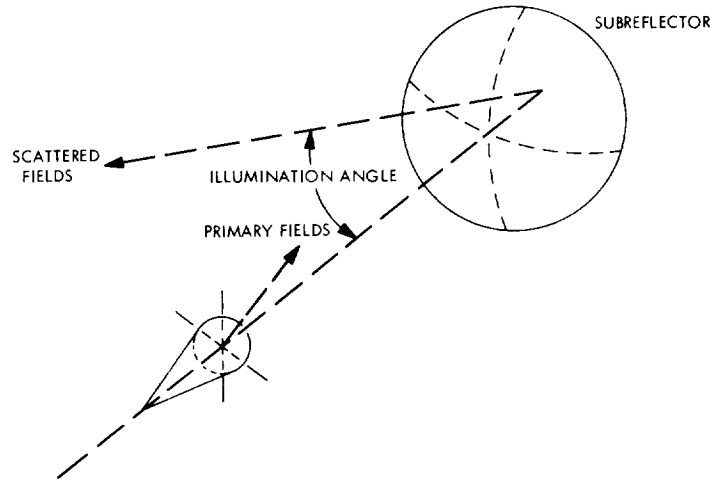


Fig. 15. Scattering configuration

Typical amplitude patterns computed by the Rusch program with two different point densities are shown in Figs. 16 and 17, and errors are given in Fig. 18. The behavior of the phase patterns is essentially identical with that of the amplitude patterns. As shown in Fig. 18, error is evaluated at illumination angles of 0, 20, and 40 deg (main beam region), 60 deg (taper region) and 80 deg (sidelobe region). The error is the absolute value of the complex difference between the computed and true patterns, expressed in decibels below the peak of the true pattern.

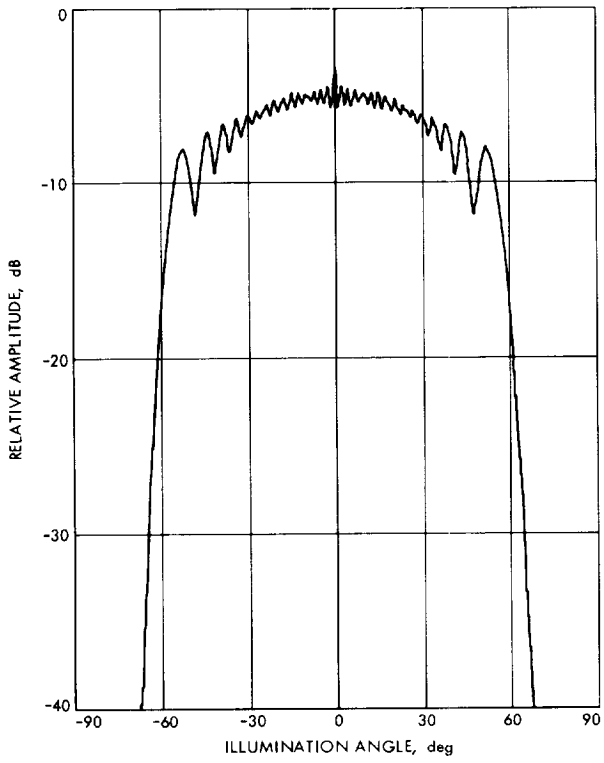


Fig. 16. Computed patterns from the Rusch scattering program, 20 points per wavelength, 0.5-deg increments

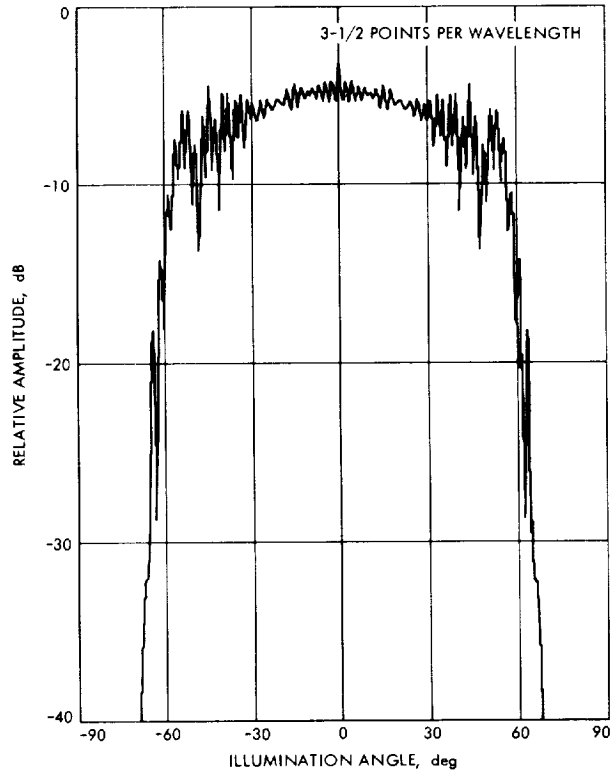


Fig. 17. Computed patterns from the Rusch scattering program, 3.5 points per wavelength, 0.5-deg increments

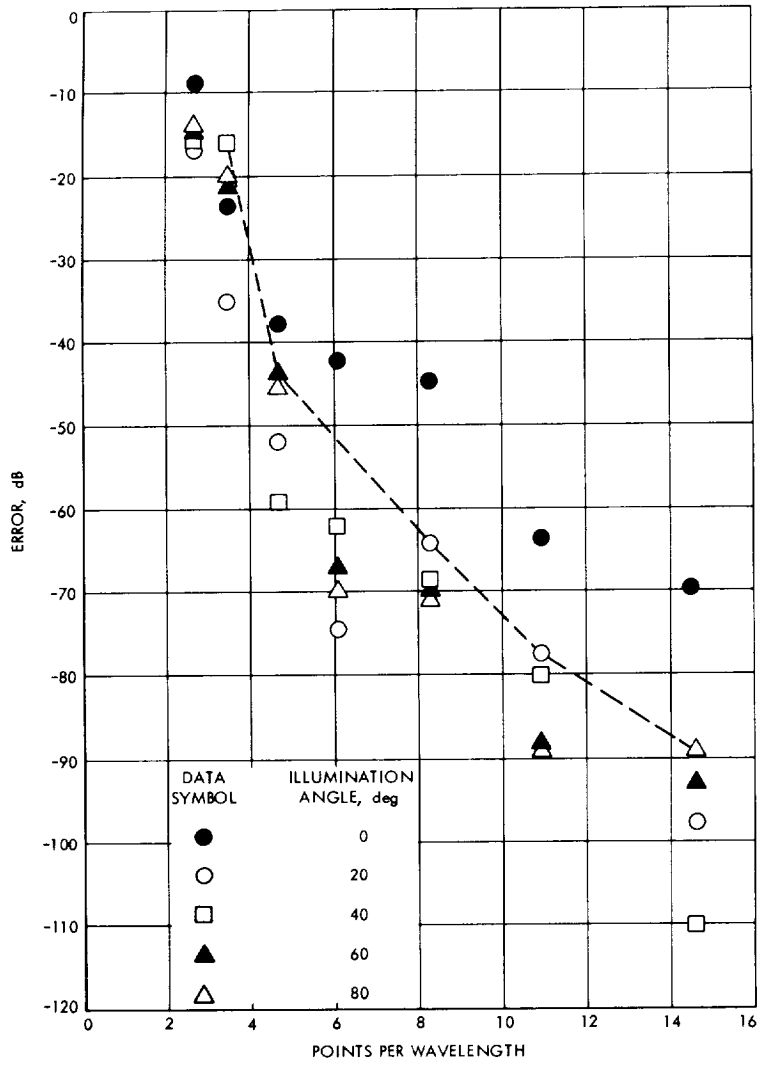


Fig. 18. Comparative error of patterns computed from the Rusch scattering program

It is seen that the error at 0 deg is particularly erratic. Since the point at 0 deg is the least important point for computing antenna performance, the error at 0 deg will be ignored; hereafter the error will mean the worst case among the errors at 20, 40, 60, and 80 deg, as shown by the dashed line in Fig. 18. On this basis roughly 5 points per wavelength are required by the Rusch program to achieve an error level below −40 dB. This density also corresponds to the minimum density required to eliminate the noise-like variations exhibited in Fig. 17.

The point density used in evaluating the two-dimensional integral in the Ludwig program is best specified in terms of points per square wavelength. However, it is of interest to consider variations of point densities in the radial and azimuthal directions separately. As one would expect, the effects of the azimuthal and radial point densities are not independent. It was found that errors were below −40 dB for one-dimensional densities of one point per wavelength with a density of 2½ points per wavelength in the other variable. It was also found that minimum error occurred when the radial and azimuthal densities were roughly equal and that 1.2 points per wavelength in both variables gave errors of less than −40 dB.

In the Rusch program, where the azimuthal density is effectively infinite, the results cannot be extrapolated to the two-dimensional case simply by squaring the one-dimensional results. However, this does yield a lower

bound, so for the purposes of comparison we can say that Simpson's rule would require *at least* 25 points per square wavelength to achieve error levels below −40 dB. Conversely, since the method of Ludwig achieves errors below −40 dB with 1 point per wavelength in one variable and a finite density in the other variable, one would expect that in the one-dimensional case this error level could be achieved with less than 1 point per wavelength using this method. It should be noted that the computer time required for a single integration point using the Ludwig techniques is estimated to be about four times as long as for a single point using the Rusch technique.

3. Accuracy of Performance Calculations

Although computed patterns are accurate on a point-wise basis, what is actually important is the overall effect on computed antenna performance. In the usual evaluation system, the computed patterns are input to other computer programs to determine antenna gain and noise temperature (Ref. 1). Using the Rusch program with output increments of 0.5 deg for the input to the analytical programs, it was found that densities greater than 5 points per wavelength gave noise temperature errors from the standard pattern of less than 0.05°K and a gain error of less than 0.001 dB. Below five points per wavelength, accuracy deteriorates rapidly.

So far, only numerical integration data point densities have been discussed; however, computer time is nearly proportional to both integration and output point densities. The parameter set used in the past with the Rusch program has had 400 integration points (IP) and 181 output points (OP), resulting in an IBM 7094 computer running time of about 36 min.

In order to determine the number of subreflector scattered pattern points necessary to accurately determine overall antenna performance, tests were run for various values of OP, with IP initially held at 15 points per wavelength (IP = 400). A reference pattern was computed for angle increments ($\Delta\theta$) of 0.25 deg. Past experience indicated that this would be more than sufficient, and the amount of predicted performance divergence as $\Delta\theta$ is increased would show whether this is a valid assumption. Other patterns were then generated for larger incremental angles, using the same number of integration points. The patterns from 0.25 to 1.0-deg increments were visually indistinguishable, but there was a noticeable loss of detail in the 2.0-deg pattern. Obviously the detail loss was worse for larger incremental

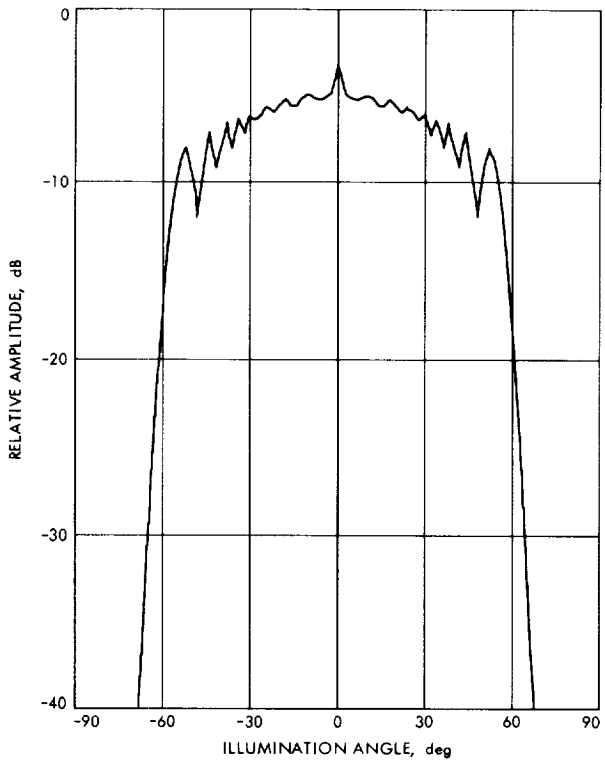


Fig. 19. AAS subreflector pattern, 15 points per wavelength, 2-deg increments

angles. Figure 19 shows the pattern for 2.0-deg increments. Note the difference between this and Fig. 16.

The patterns thus generated were input to the Ludwig antenna efficiency program (Ref. 1) and the efficiency of a parabolic antenna with an edge half-angle of 60 deg was computed for each incremental angle. Results were obtained for the overall antenna efficiency η_o, the illumination efficiency η_i, the phase efficiency η_p, and the dish edge spillover efficiency η_s. The cross-polarization efficiency was ignored because it is generally a negligible factor in a well-designed antenna.

Figure 20 shows the errors in the computed efficiencies, using the 0.25-deg increment pattern as a reference. The errors are negligible up to an increment of 1.5 deg, but beyond that point they increase rapidly and also fluctuate in value. Note for the cases studied that the errors in the individual efficiencies tend to compensate. There is no reason to conclude that this would always be true; therefore, the magnitudes of the individual errors should determine the increment size. On this basis an increment of 1.5 deg appears to be the break point, with 1.0 deg giving a comfortable margin.

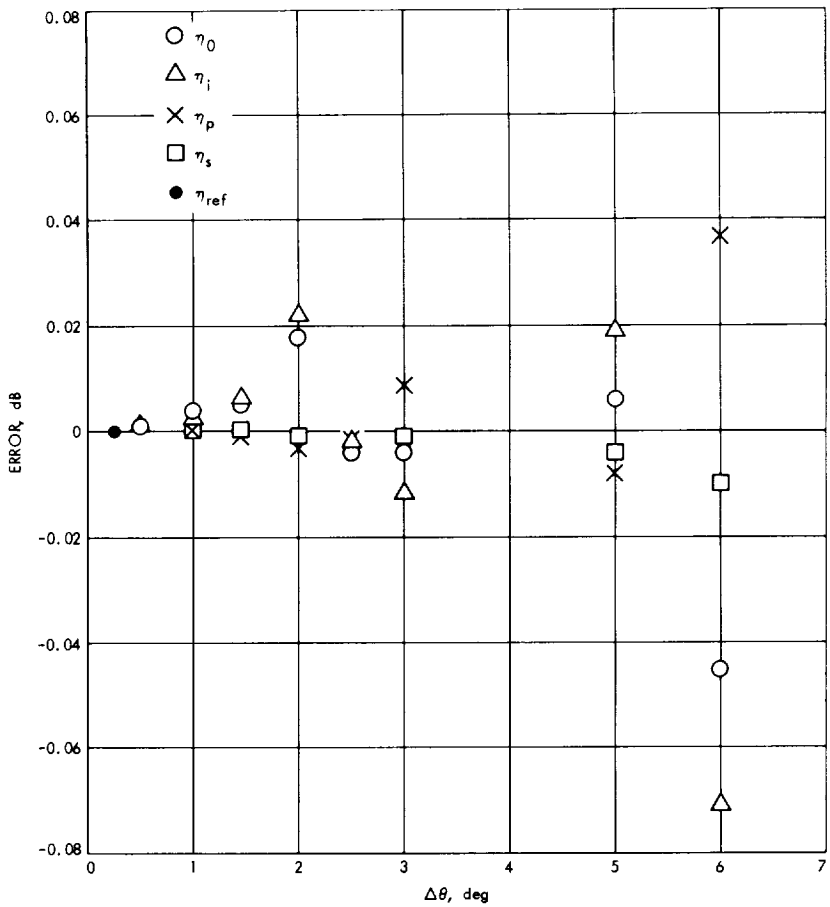


Fig. 20. Antenna efficiency error versus increment size

To test this hypothesis in conjunction with the previously developed integration point criteria, a pattern was generated using 5 points per wavelength (IP = 132) with 1.0-deg incremental angles (OP = 91). The pattern thus generated was virtually indistinguishable from the standard pattern. The computed efficiencies in decimal fraction form were identical to four significant figures ($\simeq$0.003 dB) with their 1.0-deg (IP = 400) counterparts. Thus it appears that for the Rusch program, 5 points per wavelength and 1.0-deg increments are valid criteria to use in analyzing antenna performance for the cases studied. The computer time required to run this case was about 6 min, compared to 36 min for the previously generally used case. Since, for the same case a 5-point-per-wavelength Rusch pattern and a 1½-point-per-square-wavelength Ludwig pattern are essentially identical, it can be assumed the 1-deg criteria applies to both.

The incremental angle analysis was not conducted with the generality of the integration point analysis, and the conclusions are not necessarily applicable to any case. Two observations, however, seem to point the way to a more general application. As was previously noted, the 2.0-deg pattern has a visually obvious loss of detail over the 0.25 and 1.0-deg patterns. Thus it appears that, as with the integration points, visual pattern comparison can determine if enough pattern angle points were used. The second observation concerns the ripples in the pattern. The peak-to-peak spacing of the major ripples in the amplitude and phase patterns varies between 4 and 8 deg. If this value is considered a form of "angular wavelength," then the 1.0-deg increment gives a 4 to 8 point-per-wavelength comparison, which roughly agrees with the points-per-wavelength criteria developed in the integration point analysis.

4. Other Criteria for Judging Pattern Accuracy

Preceding sections have relied on the most common method for judging the accuracy of a numerical technique; that is, to increase the number of data points to

see if anything changes. Although this is almost always a reliable method, it is possible for data to converge to an incorrect result.

One reason for accepting a pattern computed with high integration and output point densities as a true pattern is that it has been established that these patterns agree very well with experimental data (Ref. 3). Agreement between the Rusch program and a program based on a boundary value technique has also been demonstrated (Ref. 3). These facts strongly indicate that the results are basically sound.

On a numerical basis, not only do the results from the Rusch program converge pointwise to the standard pattern (which was computed using a density of 20 points per wavelength), but the results from Ludwig's program also converge to this pattern. Although both programs have Kirchoff theory in common, they differ greatly in numerical technique, and this level of absolute agreement is gratifying.

A very valuable absolute test (as distinct from the relative tests presented previously) is to compare the total output power in the scattered pattern to the total input power in the feed pattern. For integration point densities of 6 or more points per wavelength the power values are in close agreement. For low-point densities there is substantially more power in the scattered pattern than in the feed pattern—a definite indication of error.

5. Conclusions

It has been determined that for one-dimensional integrals an input data spacing of 5 points per surface lineal wavelength and Simpson's rule integration result in pointwise errors of more than −40 dB below the pattern maxima. Using the integration technique of Ludwig the spacing can be increased to about 1 point per wavelength. For two-dimensional integrals the Ludwig technique requires 1½ points per square wavelength, and Simpson's rule requires at least 25 points per square wavelength. Assuming that the Ludwig integration technique requires four times as much computer time per integration point, the two techniques require about the same computer time for one-dimensional integration and an equal number of output points. However, for asymmetrical scattering where two-dimensional integration is a requirement, Simpson's rule would require about four times as much computer time as the Ludwig technique for the same number of output points.

If output data points are spaced such that there are 4 to 8 data points for each cycle of the periodic type of variation characteristic in this type of pattern, and the integration points are spaced as described above, performance data is accurate to better than 0.01 dB.

Although this data was developed for a very particular type of scattered pattern, it is considered applicable to a fairly general class of problems, the primary exception being high-gain narrow-beam patterns. However, the numbers given above do not include much of a safety margin, and when analyzing configurations that differ markedly from the one presented here, a similar convergence test should precede detailed analysis. It was found that a visual check of computed patterns is an excellent indicator of bad data, and a reliable numerical check is a comparison of total input and output power.

The cost of the safety margin used prior to this study was shown to be about a factor of six, so although working without comfortable margins is a little troublesome, it can result in a very substantial reduction in computer time.

References

1. Ludwig, A., *Computer Programs for Antenna Feed System Design and Analysis*, Technical Report No. 32-979, Vol. I. Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1967.
2. Rusch, W. V. T., *Scattering of a Spherical Wave by an Arbitrary Truncated Surface of Revolution*, Technical Report No. 32-434. Jet Propulsion Laboratory, Pasadena, Calif., May 27, 1963.
3. Ludwig, A. C., and Rusch, W. V. T., *Digital Computer Analysis and Design of a Subreflector of Complex Shape*, Technical Report No. 32-1190. Jet Propulsion Laboratory, Pasadena, Calif., November 15, 1967.

XXI. Spacecraft Telemetry and Command

TELECOMMUNICATIONS

A. Frequency Acquisition in an MFSK Receiver,

H. Chadwick

1. Introduction

The m-ary noncoherent frequency shift-keyed (MFSK) communication link is under study as a technique for low data rate transmission at low signal-to-noise ratios (SPS 37-33, Vol. III, pp. 103–107, and Refs. 1 and 2). A possible application of this technique is in the transmission of data from a capsule landed on the surface of a planet. The available power is restricted by the size and weight of the capsule; therefore, the received signal-to-noise ratio on earth may be expected to be small. Under such conditions, the more conventional phase-locked loop receivers are not practical, whereas the noncoherent MFSK approach offers a potential solution.

There are two synchronization processes involved in an MFSK system: (1) time synchronization (i.e., the correct timing between the transmitter and receiver of the beginning of each transmitted word), and (2) frequency synchronization (i.e., the correct location at the receiver of the position of the carrier frequency in the spectrum). Inaccuracy in either of these synchronization processes results in loss of performance in an MFSK system.

Both time and frequency synchronization can be regarded as a combination of an initial-acquisition procedure, when transmission is first started, and a subsequent tracking procedure that follows any changes in the parameters during transmission. The problem of time-synchronization acquisition has been treated in SPS 37-48, Vol. III, pp. 252–264. This article deals with the problem of frequency acquisition, using some of the results in the time synchronization paper, and will show that frequency-synchronization acquisition results in considerable time lost at the beginning of a transmission. It will also be shown that the maximum relative frequency error between transmitter and receiver must be held small if frequency acquisition is to occur in a reasonable amount of time.

2. Frequency-Acquisition Problem

In the MFSK system under consideration, one of M messages, x_ℓ; ($\ell = 1, 2, \cdots, M$) is transmitted during each T-sec interval as a sinusoidal tone at a frequency f_ℓ. This tone is then modulated onto a carrier frequency, f_c, to produce the single transmitted frequency $f_c + f_\ell$. Ideally, at the receiver, the carrier frequency f_c is removed and the resulting signal, at frequency f_ℓ, is detected using

spectral analysis. However, due to instabilities in the transmitter and receiver oscillators, and uncompensated doppler shifts, it is not possible to know in advance exactly the received carrier frequency at the receiver. The actual received carrier frequency may be represented as

$$f_c + \lambda f_c = (1 + \lambda) f_c$$

where λ is a factor representing the difference between the actual and nominal carrier frequencies. The demodulation process then produces the frequency $f_l + \lambda f_c$, which effectively shifts the position of the signal in the spectrum and can cause an error in the decision process in the receiver.

The frequency shift, λf_c, must be estimated from the received signal so that the displacement of the spectrum may be effectively eliminated by changing the frequency of the local oscillator to compensate. The effects of a residual frequency error on the probability of error in the receiver are discussed in the remainder of this article. Also, several techniques for initial frequency acquisition, and their effect on system performance, are described.

3. Probability of Error Due to Frequency Uncertainty—Derivation

After carrier demodulation, the received signal is assumed to be of the form

$$x(t) = A \cos (2\pi f_s t + \phi) + n(t) \tag{1}$$

where

$f_s = f_l + \lambda f_c$ (the signalling frequency plus frequency error)

ϕ = unknown phase term

$n(t)$ = white-gaussian noise process

One form of the optimum MFSK receiver uses the discrete Fourier transform. The process $x(t)$ is low-pass filtered to a bandwidth W and sampled at a rate $2W$. If N samples are taken during the interval T ($N = 2WT$) by the sampling theorem, no information has been lost about the filtered waveform. The discrete Fourier-transform receiver calculates the quantities

$$a_k = \sum_{i=0}^{N-1} x_i \cos\left(2\pi \frac{ik}{N}\right), \qquad k = 0, 1, 2, \cdots, \frac{N}{2} \tag{2}$$

and

$$b_k = \sum_{i=0}^{N-1} x_i \sin\left(2\pi \frac{ik}{N}\right), \qquad k = 0, 1, 2, \cdots, \frac{N}{2} \tag{3}$$

where $x_i = x(iT/N)$ are the sampled values of $x(t)$.

These discrete Fourier coefficients are spaced at frequencies $f_k = k/T$ and cover the spectrum from 0 to W. If M of these frequencies are chosen as the signalling frequencies, the spacing between them will be W/M in frequency, or $N/2M$ in terms of the number of spectral lines.

A theorem attributed to Cochran, *et al.* (Ref. 3) states that given a bandlimited waveform $g(t)$, with continuous Fourier transform $G(f)$, the discrete Fourier transform of the samples of $g(t)$ at frequency k/T is $(N/T)G(k/T)$. This theorem will be used to determine the discrete Fourier transform of the received samples, x_i. Since the noise term in x_i is assumed independent of a frequency shift in the signal, only the signal term is now considered; the noise term is combined later using the results of the time-synchronization study.

The waveform $g(t)$ is a T-sec segment of a sinusoidal signal at frequency $f_s = f_l + \lambda f_c$ and arbitrary phase, ϕ. The continuous Fourier transform of such a signal is

$$\begin{aligned} G(f) = {} & \frac{AT}{2} \cos\phi \exp\left[-j(2\pi fT/2)\right] \\ & \times \left[\operatorname{sinc} 2\pi (f - f_s)\frac{T}{2} + \operatorname{sinc} 2\pi (f + f_s)\frac{T}{2}\right] \\ & + j\frac{AT}{2} \sin\phi \exp\left[-j(2\pi fT/2)\right] \\ & \times \left[\operatorname{sinc} 2\pi (f - f_s)\frac{T}{2} - \operatorname{sinc} 2\pi (f + f_s)\frac{T}{2}\right] \end{aligned} \tag{4}$$

where $\operatorname{sinc} x = (\sin x)/x$ and $j^2 = -1$.

Since the region of interest in the frequency domain is $f \approx f_s$, the terms involving $\operatorname{sinc} 2\pi (f + f_s)\, T/2$ are small and, for the purpose of simplicity, may be ignored. Under this assumption, then

$$\begin{aligned} G(f) = {} & \frac{AT}{2} \exp\left[-j(2\pi fT/2)\right] \operatorname{sinc} 2\pi (f - f_s)\frac{T}{2} \\ & \times \left[\cos\phi + j \sin\phi\right] \end{aligned} \tag{5}$$

and

$$|G(f)|^2 = \frac{A^2T^2}{4}\,\mathrm{sinc}^2\,2\pi(f - f_s)\,\frac{T}{2} \tag{6}$$

By the application of the theorem for discrete transforms

$$r_k^2 = a_k^2 + b_k^2 = \frac{N^2}{T^2}\left|G\left(\frac{k}{T}\right)\right|^2$$
$$= \frac{A^2N^2}{4}\,\mathrm{sinc}^2\left[2\pi\left(\frac{k}{T} - f_s\right)\frac{T}{2}\right] \tag{7}$$

When $f_s = l/T$, l an integer (no frequency error)

$$r_k^2 = \begin{cases} \dfrac{A^2N^2}{4} & l = k \\ 0 & l \neq k \end{cases}$$

which agrees with the results obtained previously (SPS 37-48, Vol. III). When

$$f_s = \frac{l}{T} + \lambda f_c = \frac{l}{T} + \frac{\rho}{T}$$

$$r_k^2 = \frac{A^2N^2}{2}\,\mathrm{sinc}^2\,[\pi(k - l - \rho)] \tag{8}$$

and if i is defined as $k - l$, the number of spectral lines between the nominal signalling frequency and the observed frequency, then

$$r_k^2 = \frac{A^2N^2}{4}\,\mathrm{sinc}^2\,[\pi(i - \rho)] \tag{9}$$

Remembering that the value in Eq. (9) is the magnitude of the discrete Fourier component when no noise is present, the effect of the noise may be included by analogy with the results obtained in the time-synchronization study. In the case of perfect frequency and time synchronization, the probability density of the spectral component r_k is given by (SPS 37-48, Vol. III)

$$p(r_k) = \frac{r_k}{\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\left[r_k^2 + B^2\right]\right\} I_o\left[\frac{r_kB}{\sigma^2}\right] \tag{10}$$

where $B = AN/2$ and σ^2 is the variance of the noise in either the sine or cosine transforms a_k or b_k. The value B in this expression was the contribution due to the signal, which for the case of frequency error, becomes

$$B\,|\,\mathrm{sinc}\,\pi(i - \rho)\,|$$

so that

$$p(r_k) = \frac{r_k}{\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\left[r_k^2 + B^2\,\mathrm{sinc}^2\,\pi(i - \rho)\right]\right\}$$
$$\times I_o\left[\frac{r_kB\,|\,\mathrm{sinc}\,\pi(i - \rho)\,|}{\sigma^2}\right] \tag{11}$$

when frequency error is included. The probability of error is thus given by the probability that for $k = l$ (signalling frequency l/T, $r_l > r_k$ for all other K, and is given by the expression[1]

$$P_e = 1 - \int_0^\infty z\exp\left\{-\frac{1}{2}\left[z^2 + \left(\frac{E}{No}\right)\mathrm{sinc}^2\,\pi\rho\right]\right\} I_o\left[\left(\frac{E}{No}\right)^{1/2} z\,|\,\mathrm{sinc}\,\pi\rho\,|\right]$$
$$\prod_{\substack{j=-\frac{M}{2} \\ j\neq 0}}^{\frac{M}{2}-1}\left\{1 - Q\left[\left(\frac{E}{No}\right)^{1/2}\,\left|\mathrm{sinc}\,\pi\left(\frac{jN}{2M} - \rho\right)\right|, z\right]\right\} dz \tag{12}$$

For $\lambda f_c > W/2M$, a frequency error greater than half of the spacing between signalling frequencies, the probability of a correct decision becomes small, and the system is useless. Therefore, it is only necessary to calculate the probability of error in the region $-W/2M < \lambda f_c < W/2M$ and to rely on a synchronization scheme to ensure that the frequency error will be within either this region or the limits $-N/4M < \rho < N/4M$.

To determine the effect of dropping the terms involving $f + f_s$ in the expression for r_k^2 [Eq. (9)], a computer program was run that calculated the discrete Fourier transform of a sinusoidal signal with $\rho = 0.5$. In Fig. 1,

[1]For a complete derivation of this expression, see H. Chadwick, *Frequency Acquisition in an MFSK Receiver* (JPL internal document).

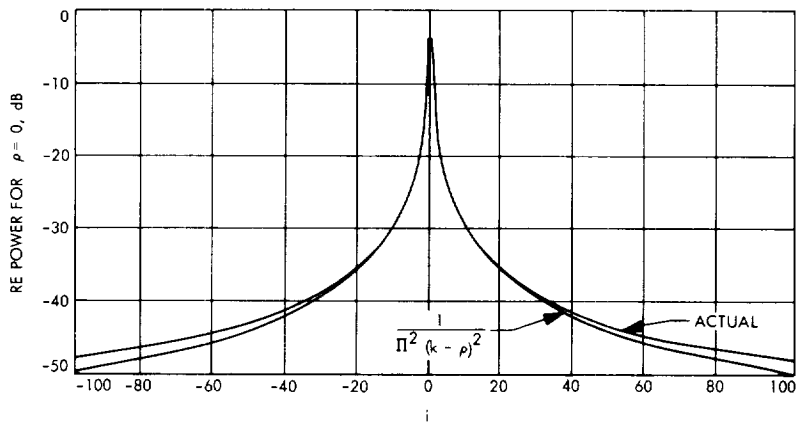


Fig. 1. Spectrum of signal at $\rho = 0.5$

the results are compared with the results obtained using the approximation.

Figures 2, 3, and 4 illustrate the calculated probability density function $p(r_k)$ for various values of ρ and i.

The probability of error due to frequency error, calculated for different values of $N/2M$, is plotted in Figs. 5 and 6. It can be seen from the curves that even with a wide signal spacing (Fig. 6), the probability of error is very sensitive to frequency variations of magnitude less than the spectral line spacing.

4. Frequency Acquisition Techniques

a. Discussion of techniques. It is evident that when an MFSK system is first turned on, the initial frequency error may be several times greater than the bandwidth W of the Fourier spectrum process. It will be assumed here that an upper bound is known to this initial error. This upper bound will be called $K_{max}W$ so that

$$-K_{max}W < \lambda f_c < K_{max}W$$

In the time-synchronization study, it was assumed that frequency synchronization had been obtained, and that

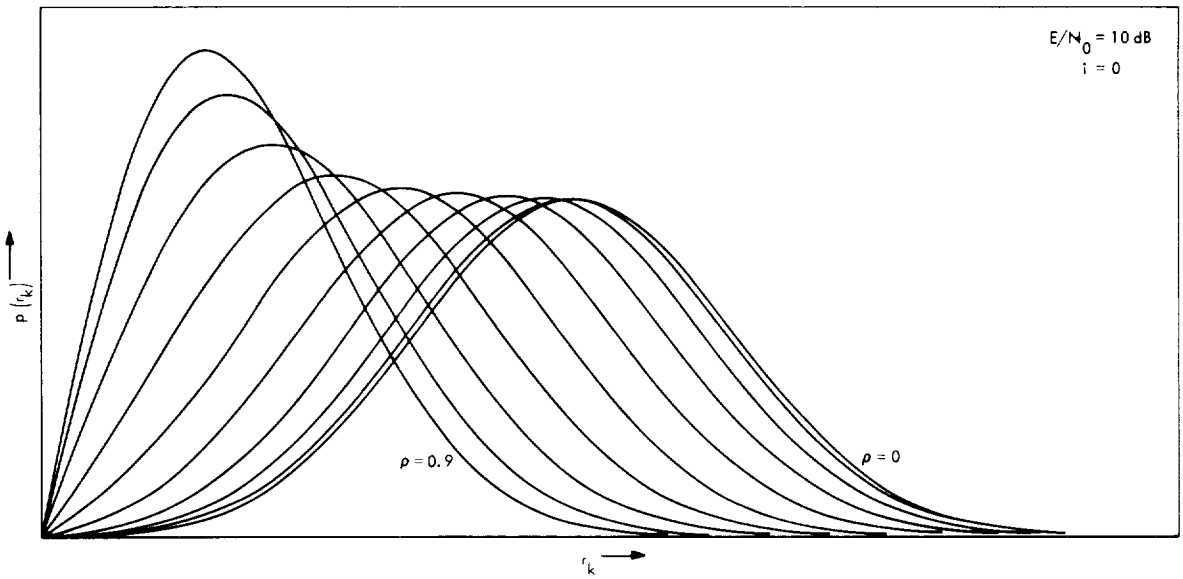


Fig. 2. Probability density function of r_k for $i = 0$

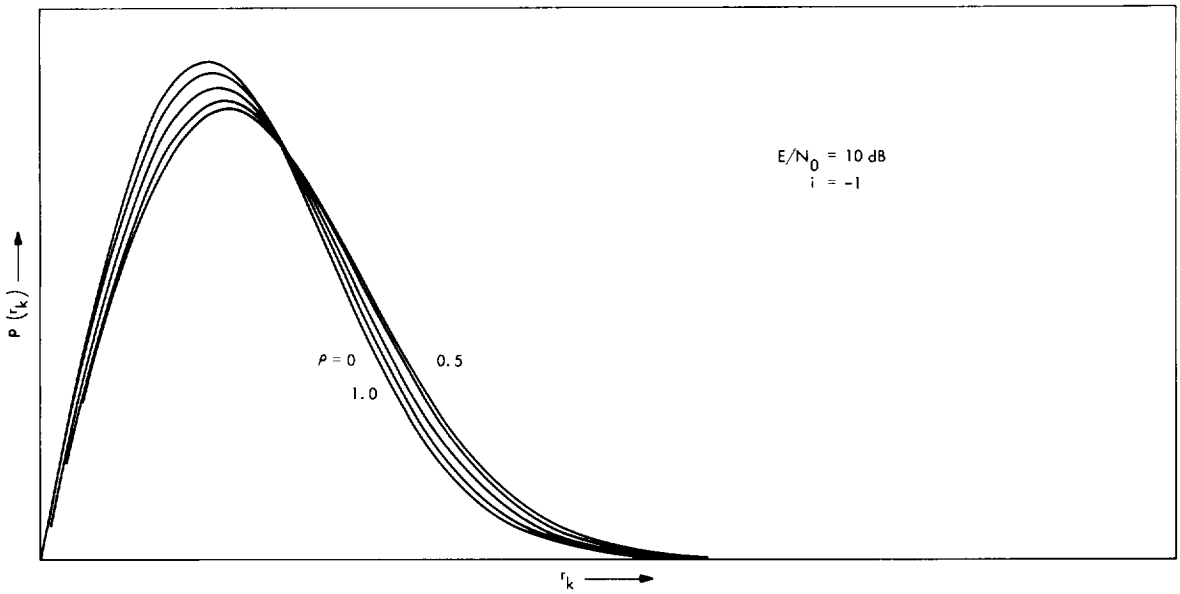


Fig. 3. Probability density function of r_k for $i = -1$

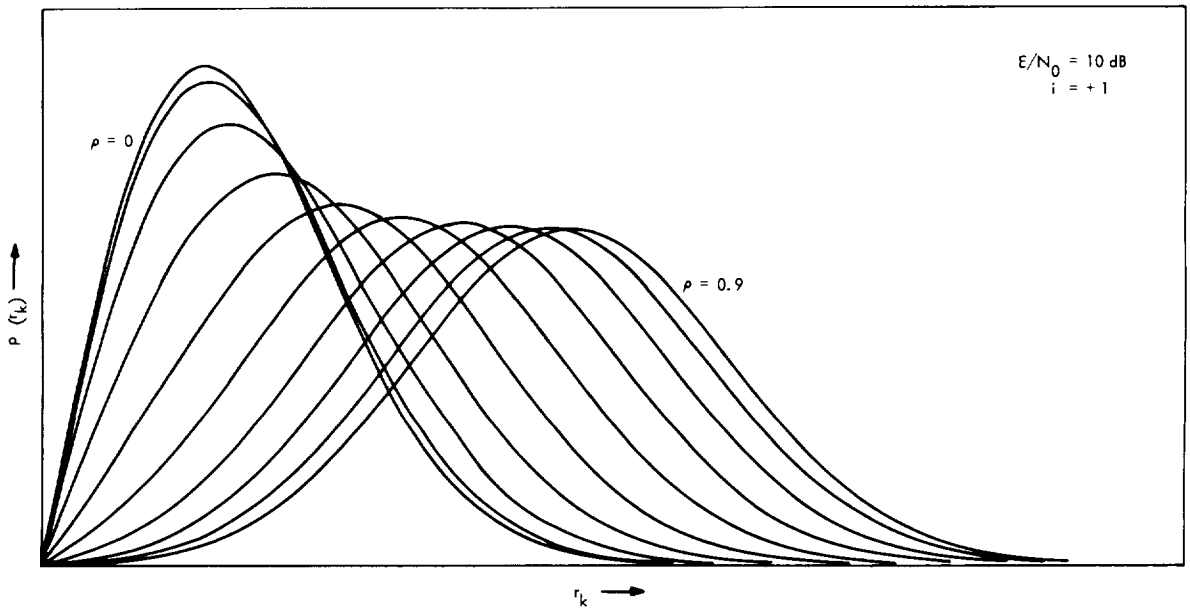


Fig. 4. Probability density function of r_k for $i = +1$

a synchronizing sequence of two alternating frequencies would be transmitted ahead of the data sequence. For frequency synchronization, therefore, it is necessary to assume that no time synchronization exists, and that the same synchronizing sequence may be used.

Probably the most straightforward frequency synchronization technique would be to step up the bandwidth of the receiver to $2K_{max}W$, sample at twice this frequency for $2T$ sec, compute the spectrum, and pick out the two signal frequency terms. This technique, however, would overburden almost any computer due to the number of samples that would have to be stored.

A more practical solution is to step the local oscillator at the receiver in frequency from $f_c - K_{max}W$ to $f_c + K_{max}W$,

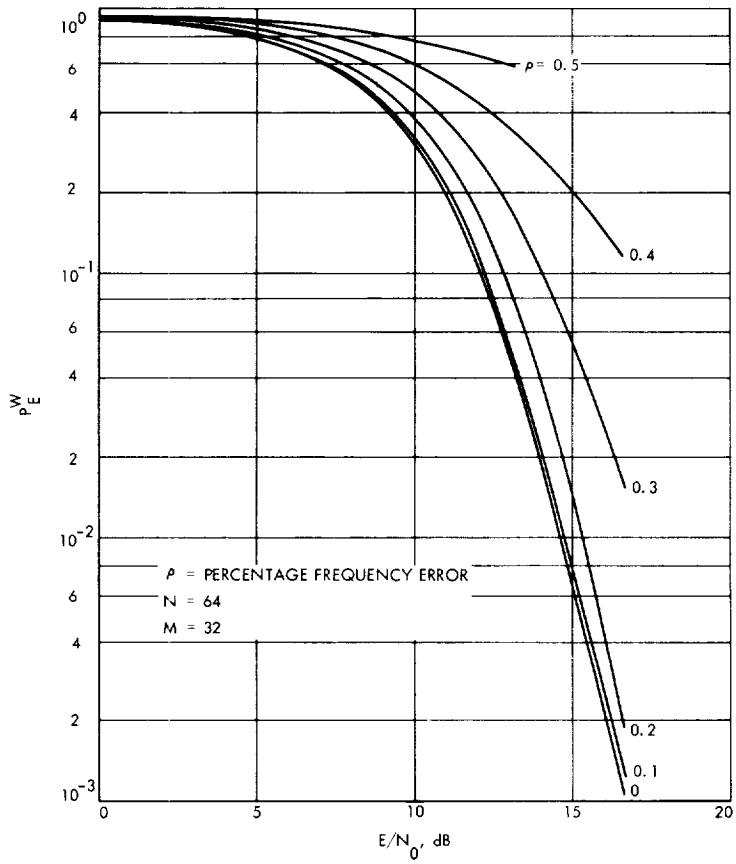


Fig. 5. Probability of word error vs signal-to-noise ratio

in steps of W, and compute the spectrum in each band of width W looking for the two signal peaks over a $2T$-sec interval. (The $2T$-sec interval is required to assure that both signal frequencies are present in equal energy.) The search may be performed either over the entire range, in which case the two largest peaks in the spectrum are chosen, or by establishing a threshold and choosing the first peak to exceed that threshold. The second technique was chosen here because it reduces the average time necessary to perform acquisition.

b. Derivation of false-acquisition and no-acquisition probabilities. Because of the low expected signal-to-noise ratios at the MFSK receiver, it will probably be necessary to average the spectrum in each band over several $2T$-sec intervals. The frequency error, λf_c, can be written

$$\lambda f_c = KW + \frac{k}{T} + \frac{\rho}{T}$$

where

$$|K| = 0,1,2,\cdots,K_{\max}$$

$$k = 0,1,2,\cdots,\frac{N}{2} - 1$$

$$0 \leq \rho < 1$$

and with the two signalling frequencies $f_L = l_L/T$ and $f_H = l_H/T$, the spectral lines due to the signals will appear in the Kth band at frequencies

$$\frac{k}{T} + \frac{\rho}{T} + \frac{l_L}{T}$$

and

$$\frac{k}{T} + \frac{\rho}{T} + \frac{l_H}{T}$$

where $l_H > l_L$.

The probability density of the magnitude of the Fourier component nearest to each of the signalling frequencies, at $(k + l_L)/T$ and $(k + l_H)/T$, will depend on the value of ρ. The probability density function of this component

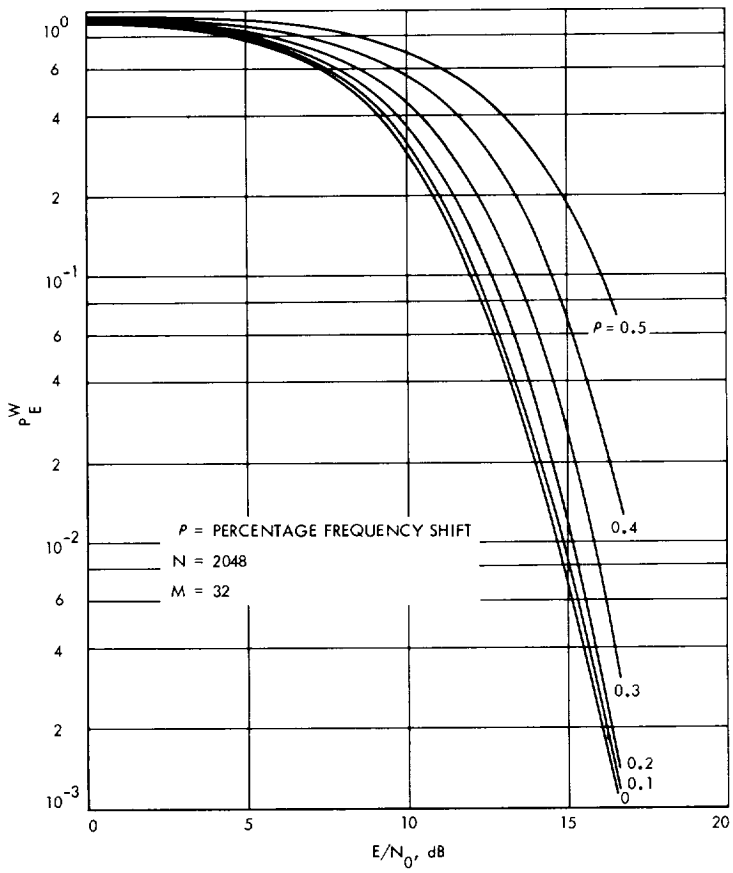


Fig. 6. Probability of word error vs signal-to-noise ratio

is

$$p(r_k) = \frac{r_k}{\sigma^2} \exp\left[-\frac{1}{2}\left(\frac{r_k^2}{\sigma^2} + \gamma^2\right)\right] I_o\left(\frac{\gamma r_k}{\sigma}\right) \qquad (13)$$

where γ^2 is the effective signal-to-noise ratio given by

$$\gamma^2 = \frac{n}{2}\left(\frac{E}{No}\right) \operatorname{sinc}^2 \pi\rho \qquad (14)$$

where n is the number of times each spectrum is averaged in each band.

The lowest value of γ^2 occurs for $\rho = 0.5$, or for the signalling frequency exactly midway between spectral lines. At this value,

$$\gamma^2 = \frac{n}{2}\left(\frac{E}{No}\right) \operatorname{sinc}^2 \frac{\pi}{2} = \frac{2n}{\pi^2}\left(\frac{E}{No}\right)$$

The other spectral lines are largely due to noise. As Fig. 1 illustrates, the signal component at the second closest spectral line is 10 dB below that of the closest spectral line for $\rho = 0.5$. This component is lower for other values of ρ. To simplify the computations, it has been assumed that all spectral lines more than one away from the signal are due to noise only, with the probability density function

$$p(r_i) = \frac{r_i}{\sigma^2} \exp\left[-\frac{r_i^2}{2\sigma^2}\right] \qquad (15)$$

In the threshold system for frequency acquisition, a threshold is established based on the values of the spectral components due to noise only. This threshold, r_o, is set so that the probability of any noise component exceeding r_o is less than a predetermined value, α. The probability α is, therefore, the probability of false acquisition. Similarly, once the threshold is established, the probability β of the signal component not exceeding the threshold, or no acquisition, can be computed from the probability density function of the signal component. Since there are actually two signal components, it is assumed that in the search upward through the frequencies, the lower one will be recognized first. The possibility that the lower frequency will be below the threshold and the upper frequency above the threshold is included in the probability of no

acquisition, since only the lower frequency is used to determine this probability.

Since there are at most $2K_{max}(N/2)$ spectral lines to be searched, and the noise component is statistically independent in each one, the total probability α is approximately

$$\alpha \approx Pr\,[r_i > r_o]\,NK_{max} \tag{16}$$

where

$$Pr\,[r_i > r_o] = \int_{r_o}^{\infty} \frac{r}{\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right] dr = \exp\left[-\frac{r_o^2}{2\sigma^2}\right] \tag{17}$$

Therefore

$$r_o^2 = -2\sigma^2 \ln\left(\frac{\alpha}{NK_{max}}\right) \tag{18}$$

If the variance, σ^2, is not known, it may be estimated by the maximum likelihood estimator

$$\hat{\sigma}^2 = \frac{1}{2N} \sum_{i=1}^{N} r_i^2 \tag{19}$$

With the threshold determined, the probability β may be determined by the relation

$$\beta = P_i\,[r_k < r_o] = \int_o^{r_o} \frac{r_k}{\sigma^2} \exp\left[-\frac{1}{2}\left(\frac{r_k^2}{\sigma^2} + \gamma^2\right)\right] I_o\left(\frac{\gamma v k}{\sigma}\right) dr_k = 1 - Q\left(\gamma, \frac{r_o}{\sigma}\right) \tag{20}$$

where $Q(\ )$ is Marcum's Q Function (Ref. 4). This equation can, in turn, be used to determine the requisite n for a given probability β.

c. Average time to acquisition. With no prior knowledge of the distribution of the initial frequency error, the "worst case" assumption is that of a uniform probability distribution from $-K_{max}W$ to $+K_{max}W$. For this assumption, the average time to acquire would be one half of the time to examine the entire spectrum, or

$$T_{acq} = \frac{1}{2}(2T)(n)(2K_{max}) = 2nTK_{max} \tag{21}$$

since each band is searched for $2nT$ sec and there are $2K_{max}$ bands to search.

A more reasonable assumption would probably be that the initial frequency error is centered at the frequency f_c with a smaller variance than that of the uniform distribution. For this case, the optimum search procedure would be to search the most probable bands (those close to f_c) first and then proceed to the less probable bands, thus lowering the average acquisition time. An example of this type of procedure is given in *Subsection 4-d.*

d. Computation of acquisition time for typical example. In the following computation, values have been chosen for a typical MFSK receiver which seem reasonable in the present state of knowledge. These values are[2]

$f_c = 2 \times 10^9$ Hz

$\lambda_{max} = 10^{-6}$

$W = 102.4$ Hz

$T = 10$ s; $\frac{1}{T} = 0.1$ Hz = spectral line spacing

$N = 2048$

$M = 32$

$N/2M = 32$; $W/M = 3.2$ Hz = signal frequency spacing

$K_{max} = 20$

$\alpha = \beta = 0.01$

$P_E^w = 0.01$

Using these values in the above equations, the required effective signal-to-noise ratio, γ^2, is

$$\gamma^2 = 60.0 = 17.8 \text{ dB}$$

From Eq. (15)

$$\gamma^2 = \frac{2n}{\pi^2}\left(\frac{E}{No}\right)$$

at the worst case value of $\rho = 0.5$ or

$$10 \log n = 24.7 - \left(\frac{E}{No}\right) \text{dB}$$

[2]These values are based on those used by Ferguson (Ref. 3) and by Boyd, D., *Some Practical Aspects of the Design of MFSK Telemetry Systems*, Feb. 2, 1967 (JPL internal document).

For an E/No of 16 dB, which would yield a P_e^w of about 0.01 in a synchronized system, n is approximately 7. Therefore, each band spectrum must be averaged 7 times to achieve a good probability of correct acquisition.

In terms of time, again using the same typical values, the average time to acquisition for an uniformly distributed frequency error, given by Eq. (21), is

$$T_{acq} = 2nTK_{max} = 2800\text{ s} = 47\text{ min}$$

Using a gaussian distribution for the initial frequency error, with $3\sigma = K_{max}$, and an optimized search strategy, it was found that the average number of bands searched was reduced from 20 to 11. This provides a reduction in the average acquisition time to 1540 s, or 26 min.

5. Conclusions

Obviously, the time required to obtain acquisition in the example illustrated in *Subsection 4-d* is excessive if only a short-duration data transmission is anticipated. This time, however, depends directly on the value of K_{max} (the number of bands that must be searched) which, in turn, depends on the expected relative stability of the transmitter and receiver oscillators. Any improvement in this stability will provide a linearly proportional improvement in the average synchronization time.

It is also apparent that a frequency tracking scheme is essential in addition to the initial acquisition scheme. The probability of error in the receiver is greatly affected by the stability of the signal frequency even if an averaging scheme, such as that suggested by Ferguson (Ref. 2), is used. In this article, no consideration has been given to various averaging techniques as these more properly belong to the area of frequency tracking, not initial acquisition.

References

1. Charles, F., and Springett, J., *The Statistical Properties of the Spectral Estimates Used in the Decision Process by a Spectrum Analyzer Receiver,* Paper presented at the National Telemetry Conference, San Francisco, Calif., 1967.

2. Ferguson, M., *Communication at Low Data Rates: Spectral Analysis Receivers,* Technical Memorandum 124. Philco-Ford Corporation, Space and Re-entry Systems Division, Sept. 1967.

3. Cochran, W., *et al.*, "What is the Fast Fourier Transform?" *IEEE Trans. Audio Electroacoust.*, AV-15, pp. 45–55, June 1967.

4. Marcum, J., *Table of Q Functions,* Report M-339. The Rand Corporation, Jan. 1950.

XXII. Spacecraft Radio

TELECOMMUNICATIONS

A. Low Data Rate Telemetry RF System Development, *R. Postal*

1. Introduction

A solid-state m-ary noncoherent frequency-shift keyed (MFSK) 2295 MHz transmitter has been developed as a subassembly for a telecommunication system capable of surviving a high impact on a planetary surface. A previous article (SPS 37-48, Vol. III, pp. 284–285) contained a block diagram of the transmitter. A circuit description was given in SPS 37-40, Vol. IV, pp. 198–201. An overall description of the transmitter and recent test results are given in this article.

2. Transmitter Description

Figure 1 is a photograph of the S-band, high-impact, solid-state transmitter. The transmitter circuitry and packaging for modules 0, 1 and 2 were developed at JPL. Module 3, a stripline X-4 frequency multiplier, was developed by the Motorola Corporation. A semiconductor complement of 8 transistors, 8 varactor diodes, and 2 regulating diodes is used in twelve RF stages to provide a 2295-MHz transmitter output level of 5 W. For high impact survival, all components are bonded to the circuit boards and chassis with solethane. The coil forms, shown in Fig. 1 (module 2), are an integral part of the circuit board and rigidly confine the inductors during shock environment.

3. Test Results

The transmitter develops a 2295-MHz level of 5 W at 25°C with a 13.5% dc-to-RF conversion efficiency. Figure 2 shows output level and conversion efficiency as a function of temperature. Coherent sidebands are greater than 40 dB below the output carrier; spurious sidebands are greater than 60 dB below the carrier. Performance of the transmitter has been observed over a wide range of load variations. No breakup in spectrum was noted until the load voltage standing-wave ratio (VSWR) was increased to 5:1. At this VSWR, 15 MHz sidebands were observed (only under a critical setting of load phase, however). Sensitivity of the MFSK modulator is 28 Hz/V (measured at the transmitter output). The transfer curve deviates less than 2% from a linear characteristic over a maximum carrier excursion of ± 170 Hz. Figure 3 shows a plot of transmitter frequency drift during a 20-min operating interval at ambient temperature. As shown, a majority of the drift occurred during the first 2 min after transmitter turn-on. Phase jitter at 2295 MHz is 5 deg

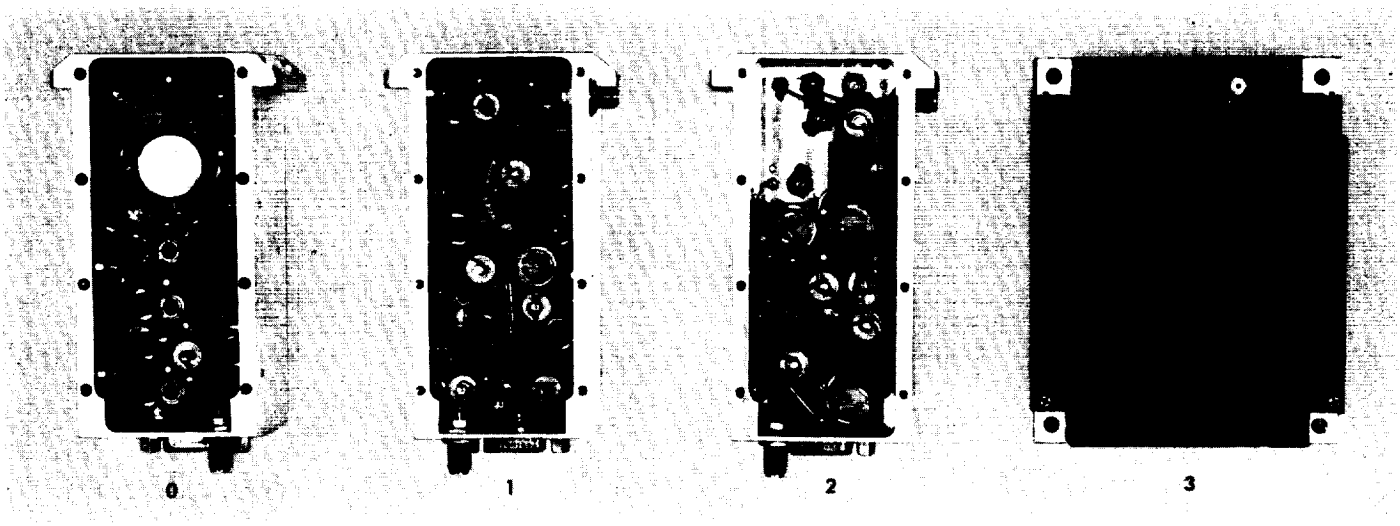


Fig. 1. Solid-state transmitter

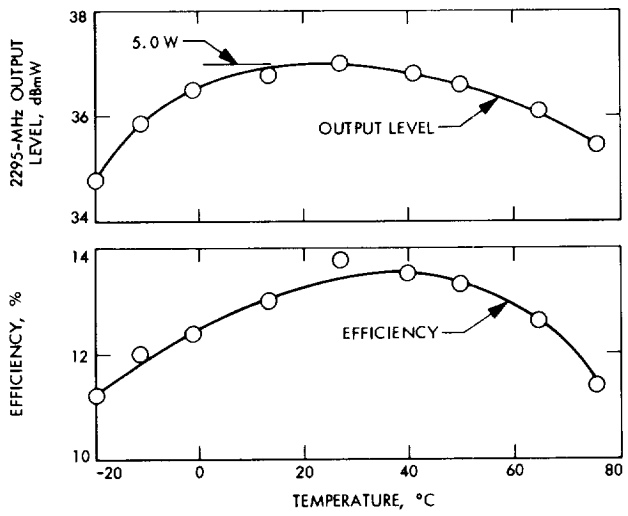


Fig. 2. Transmitter output level and efficiency as a function of temperature

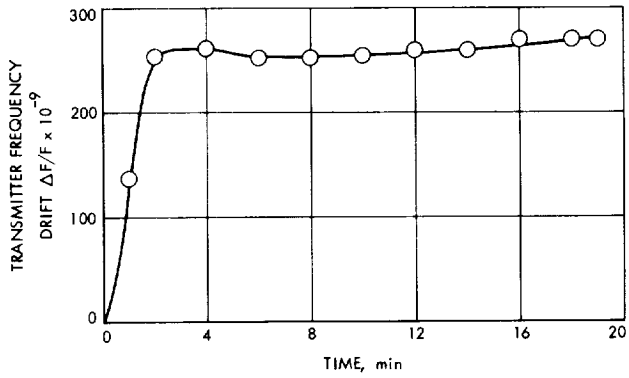


Fig. 3. Transmitter frequency drift vs time

peak as measured on a phase-locked loop receiver with a 12 Hz noise bandwidth.

The circuits and components of modules 0, 1 and 2 have survived shock levels of up to 10,000 g's. The complete transmitter was installed in the Capsule System Advanced Development (CSAD) Lander Feasibility Model for environmental testing at system level. The system was subjected to two shock tests, each preceded by a sterilization heat cycle of 125°C for sixteen hours in a 98% GN_2 atmosphere. The first shock level was calculated to be 1500 g's and the second shock was 2500 g's. Maximum frequency shift due to shock was 1 part/10^6. No other change in transmitter performance was noted at the conclusion of these tests.

B. RF Power Amplifier Life Test Summary,

R. Hughes

1. Introduction

This is the first of a series of articles dealing with the life-test performance of RF power amplifiers. These amplifiers are being life tested to advance the general knowledge of each device and to determine their applicability to future long-life spacecraft requirements. This article updates a previous life test report[1] and, specifically, describes performance data for a Hughes Aircraft Company traveling-wave tube (TWT), Model 216H, Raytheon amplitrons, Models QKS 1300, and Watkins Johnson TWTs, Models WJ 274-1 and WJ 274-6.

2. Hughes Aircraft Company Model 216H TWT

A 10-W Hughes Aircraft TWT was one of the RF power amplifiers used in the *Mariner* Mars 1964 and *Mariner* Venus 67 missions. A tube of this type, S/N 30, and a flight-type power supply (Engineered Magnetics Model EMPS 121A, S/N 16056) were placed on life test in the atmosphere, at room temperature, on January 22, 1965 (see Footnote 1). This TWT and power supply have now accumulated 29,100 h of operation. After 17,692 h, the RF output was only 0.4 dB below its initial value of 41.0 dBmW (see Footnote 1). The data in Fig. 4 show the performance from 17,692 h to the present time (the plots were derived from data points typically taken at 12-h intervals). It is evident from the data that only small changes in performance have occurred; the RF output at the 29,100-h mark is 40.4 dBmW, which is only 0.6 dB below its initial value of 41.0 dBmW. The base for the oxide cathode in the model 216H TWT is Cathalloy A33 ultrapure nickel with 0.1% zirconium and 2% tungsten. The cathode operates at a current density of 380 mA/cm^2.

3. Raytheon Model QKS 1300 Amplitrons

Three Raytheon amplitrons were purchased expressly for the purpose of determining their life expectancy. This cross-field amplifier is designed to operate at around 2.3 GHz with an RF output of 25 W and an overall efficiency of greater than 50%. Special life-test power supplies were purchased with the amplitrons. The supplies were designed to maintain the constant anode current and heater power necessary to maintain lock in the amplitron and obtain useful RF output power at the input frequency. In an out-of-lock condition, the amplitron is a noisy, free-running oscillator and, thus, the RF output is not coherent with the input.

The life test on the amplitrons was performed in the atmosphere and at room temperature. However, the heat sink on which the amplitrons were mounted was found to be typically 90°F. The RF input to the amplitrons was maintained at 27.8 dBmW and 2.295 GHz. The data obtained from the amplitron life tests, plotted in Figs. 5, 6, and 7, show that the three amplitrons had a life expectancy ranging from 2,350–3,275 h, and that the end of life occurred rather abruptly. At the end of life, the amplitrons could not sustain normal cathode current and ceased to exhibit gain.

During the tests, it was necessary to adjust the operating point of each amplitron. (These adjustments are represented by abrupt changes in the anode current shown in Figs. 5, 6, and 7.) The adjustments were necessary to compensate for changes in the volt–ampere characteristics of the amplitrons and bring the amplitrons back into a locked condition. Since the amplitrons have a relatively short life, and exhibit changes in their volt–ampere characteristics during their life, it is recommended that amplitron Model QKS 1300 not be used on any space programs requiring unattended operation.

4. Watkins Johnson WJ 274 TWTs

Three Watkins Johnson TWTs were purchased for evaluation and life testing. Two of the tubes were Model WJ 274-1 and the other was a high-efficiency Model WJ 274-6. These tubes are designed for a 50,000-h life expectancy and use an oxide cathode on 220 nickel. The cathode operates at a current density of 200 mA/cm^2. When operated at 2295 MHz, the Model WJ 274-1 TWTs typically exhibit 24 W of RF output at saturation, 24.5 dB gain, and an overall efficiency of 34%.[2] At the same frequency, the Model WJ 274-6 TWT produced 25 W of

[1]Hughes, R. S., *Life Test Report*, Aug. 1, 1967 (JPL internal document).

[2]Hughes, R. S., *WJ 274-1 Evaluation Report*, Oct. 1967 (JPL internal document).

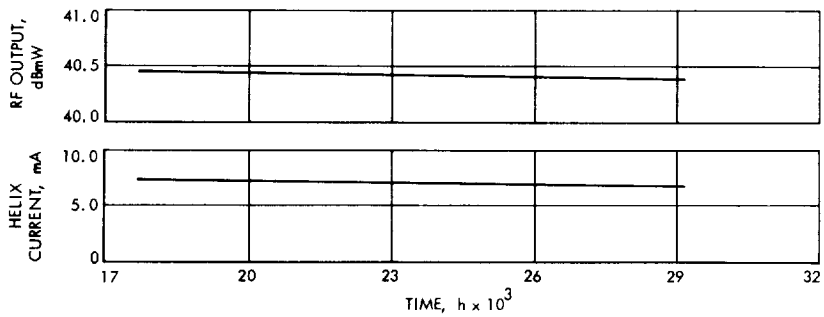


Fig. 4. TWT 216H S/N 30 life-test data

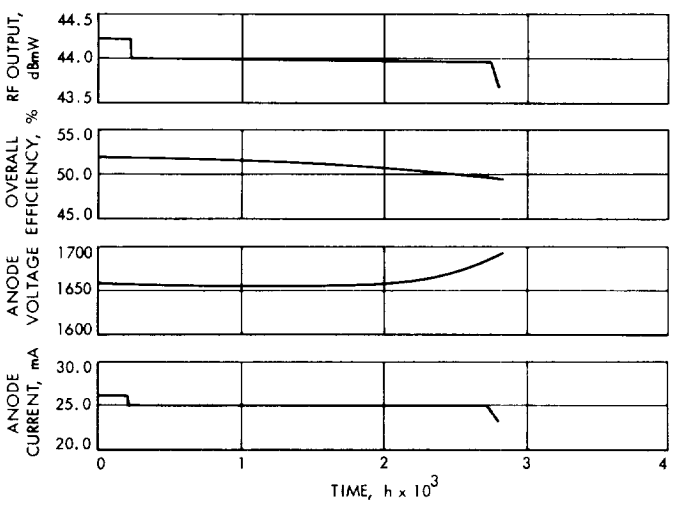


Fig. 5. Amplitron S/N 218 life-test data

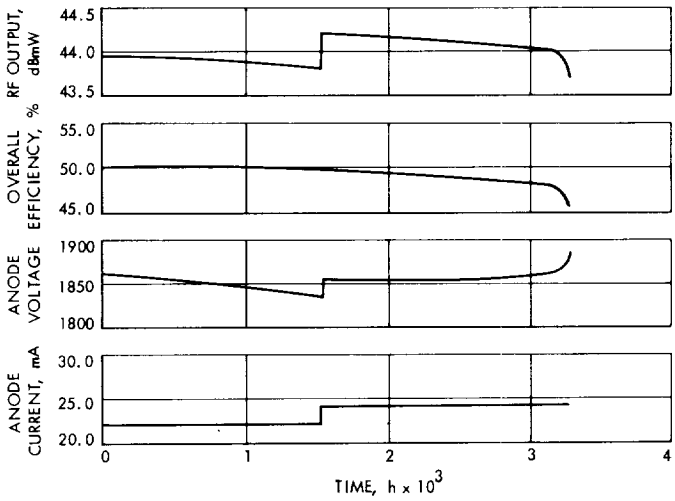


Fig. 6. Amplitron S/N 229 life-test data

RF output at saturation, 27.9 dB gain, and an overall efficiency of 41.6%.

The Model WJ 274-1 TWTs successfully passed the electrical and environmental tests (Footnote 2). The electrical tests were designed to thoroughly evaluate the TWTs performance under a variety of input-voltage, input-drive, and output-load conditions. The environmental tests consisted of the type approval static acceleration, vibration, and shock tests specified in the *Mariner* Mars 1969 Environmental Test Specification

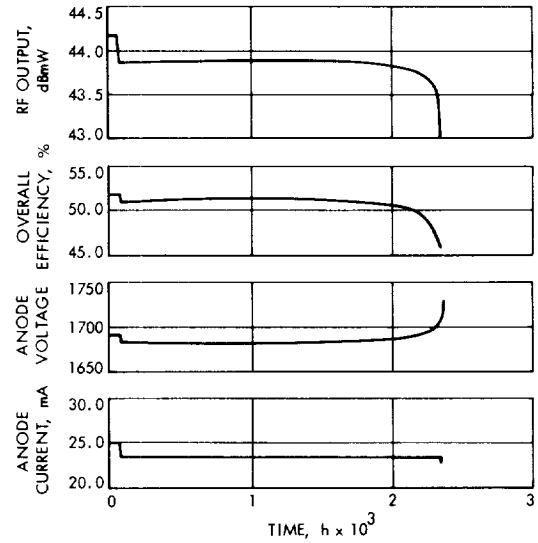


Fig. 7. Amplitron S/N 250 life-test data

TS 500437. In addition, the TWTs were subjected to a temperature test over a −10 to +75°C range. The results indicate that the Model WJ 274-1 TWT appears suitable for spacecraft applications.

The Model WJ 274-6 TWT successfully passed the battery of electrical tests listed in Footnote 2, and was found to have the same general characteristics as the Model WJ 274-1 TWT. The only environmental tests performed on the Model WJ 274-6 TWT were the type approval vibration test, specified in TS 500437, and a thermal-vacuum test over a −10 to +75°C range. The tube passed these tests without any changes in performance.

The Model WJ 274-1 TWTs were placed on life test in the atmosphere on November 9, 1967. The Model WJ 274-6 TWT is being life tested in a vacuum (its life test began on May 20, 1968). The temperature of all three TWTs is running at about 100°F near the collector. The two Model WJ 274-1 TWTs (S/N 29 and S/N 31) have each accumulated 5,050 h of operation; during this time, the RF output from each tube has changed less than 0.1 dB. In addition, there have not been any significant changes in the other parameters. The initial RF output of S/N 29 and S/N 31 was 44.1 and 44.0 dBmW, respectively. The initial RF output of the Model WJ 274-6 TWT (S/N 3) was 44.1 dBmW. This tube has been on life test for 900 h and no significant changes have occurred in its performance.

XXIII. Future Projects

ADVANCED STUDIES

A. The Objectives for Roving Vehicles in a Lunar Exploration Program, *R. G. Brereton*

The very nature of the lunar exploration task suggests that a surface mobility system will be required to acquire the needed data. Manned exploration on the lunar surface can be expected to cover only a small fraction of the moon and most of this in easily accessible spots where landings can be made safely. For short stay-time manned missions, a lunar flying vehicle seems to offer the most practical technique for allowing the geologist–astronaut to visit the most outcrops, or interesting lunar features, within the time constraints of a life support system; however, even with this technique each lunar mission will explore only a very small percentage of the moon. Obviously, unmanned surface roving vehicles must be relied upon to fill in the required detail of lunar geology and geophysics between the sparingly placed manned sites. A rover concept as a part of the lunar exploration program is required if the resources for investigating the moon are to be utilized thoroughly and economically.

The contribution of the rover on a long traverse mission to the objectives for geological exploration of the moon are shown in Table 1. In this table the essential and guiding objectives for geological exploration of the moon, as specified by the Geology Working Group of the 1967 Summer Study of Lunar Science and Exploration (Ref. 1), are shown in relation to the results for each objective that can be expected from properly instrumented traverse rovers. Generally, the contributions of the long traverse mission to geological objectives for lunar exploration are:

(1) Provide surface viewing along a traverse or of an interesting area away from the manned landing site.

(2) Provide geochemical data along a traverse or of an interesting area away from the manned landing site.

(3) Provide geophysical data along a traverse or of an interesting area away from the manned landing site.

The objectives for geological exploration of the moon (Table 1) are quite comprehensive and a variety of rover

missions and science tasks are required to provide meaningful data. Table 2 lists five possible roving vehicle missions and the appropriate scientific instruments that would form their payloads. A common rover design would serve as a bus for each mission; however, instrumentation and operating mode would vary as a function of the scientific task to be accomplished. Each of these tasks has its place in the overall lunar exploration program and so any plan that defines the most feasible and economical lunar exploration program must consider a mix of these roving vehicle tasks with other lunar missions, both manned and unmanned. These science tasks and the appropriate scientific instrumentation and operating modes for each have been discussed previously in SPS 37-51, Vol. III.

Table 3 has been prepared as a conclusion to show the consequences on the lunar exploration program of automated roving vehicles.

Reference

1. *1967 Summer Study of Lunar Science and Explorations*, NASA SP-157. National Aeronautics and Space Administration, Washington, 1967.

Table 1. Contribution of a roving vehicle to the objectives for geological exploration of the moon

Essential objectives	Expected results from a roving vehicle
To determine at appropriate scales the type, form, structure, distribution, and relative-age relations of the various masses of material which constitute the accessible portions of the moon.	An imaging system on a rover will make an order of magnitude contribution to this geomorphic and stratigraphic objective at a scale and from a viewing position unattainable from orbiters and over an area far into the environs of each post-*Apollo* site.
To determine the physical, chemical, mineralogical, and petrogenetic nature of lunar materials, both surficial and deep-seated.	The information needed for this objective requires direct lunar surface sample points from a variety of features and geographic positions. A properly instrumented rover could provide these data for the areas between post-*Apollo* landing sites thus providing a more comprehensive picture of lunar geochemistry.
To characterize from direct observations the operative processes (and their products) that are actively modifying the superficial features of the lunar surface; includes both endogenetic (volcanism, tectonism) and exogenetic (meteorite impact, solar wind) processes.	The detail of tactile visual and other information that would be available from a roving vehicle traverse of areas of the moon that may never be visited by man will add much new data to this objective.
To evaluate in the light of observational data past and current processes that may have contributed to the origins and evolution of the present major structural and lithological features of the moon, including accretion and impact of extralunar bodies, volcanism, gas emanation, deep-seated magmatism, tectonism, and other processes reflecting important mechanisms of energy transfer in the moon.	The detail of tactile visual and other information that would be available from a roving vehicle traverse of areas of the moon that may never be visited by man will add much new data to this objective.
To interpret the most probable extension with depth of major crustal features from the nature and geometry of their surface exposures and with the aid of geophysical methods.	The solution of many of the problems posed by this objective can be solved by combining the powerful depth probing techniques of traverse geophysics and the mobility of a roving vehicle.
To develop a comprehensive geological history of the moon integrating the results of both direct and remote geological investigations with geochemical and geophysical studies to establish the times of initiation, duration, and product formation for the major episodes in lunar history.	The reconnaissance type data that can be provided by a roving vehicle for the large areas between each post-*Apollo* site would contribute to the solution of this objective.
To provide a physical understanding and historical perspective for the ecological setting and protoorganic material base in which the presence or absence of lunar life is established.	The reconnaissance and search capability of a properly instrumented rover for locating protoorganic material (some precursor or fossil of life) could be very important here.

Table 2. Roving vehicle missions for geological exploration of the moon

Mission	Instrumentation	Remarks
Geochemistry (sample return)	Photo-imaging system, particulate sampler, hard rock drill, sample container (total weight, including samples: 150 lb)	Rover would collect rock samples in the environs out from a post-*Apollo* site or more probably on an extended traverse between post-*Apollo* sites. This type of mission will relate the geology of the great tracts of land between the post-*Apollo* sites to the sites themselves.
Geochemistry (*in situ* analysis)	Imaging system, X-ray spectrometer, X-ray diffractometer, petrographic microscope, particulate sampler, hard rock drill (total weight: 100 lb)	This type of mission will provide geological reconnaissance of great surface areas of the moon that cannot be investigated by man either because of time or location. Vehicle would operate for months and over traverses of hundreds of kilometers.
Traverse geophysics	Imaging system, magnetometer, gravimeter, active seismic, laser and radio ranging experiment (total weight, including 100 quarter-pound charges: 125 lb)	This mission will provide unique data toward the solution of problems that can be solved only by geophysical techniques. This is a powerful tool for providing data on the subsurface of the moon.
Deep seismic	Auger and seismic charges (total weight: 250 lb)	The rover equipped with an auger for shot-hole preparation and a number of charges would seismic-profile out from a fixed seismic station as could be provided by the *Apollo* lunar surface experiments package or the emplaced science station.
Instrument deployment	Variable	There are areas on the moon that are too rough, or inaccessible for landing vehicles because of one reason or another, and where it may be desirable to place scientific instrumentation. The rover could deliver instrument packages over hundreds of kilometers.

Table 3. The need for rovers in lunar exploration

Scientific discipline or mission	Consequences of the use of a rover
Traverse geophysics	Without a surface roving capability, this task will literally not be accomplished. This means the very powerful techniques for subsurface probing offered by active seismic gravity and magnetic surveys will find little application.
Lunar chemistry (including mineralogy, petrology, and elemental analysis)	Without a surface roving capability this task will be accomplished for a very limited area around a few landing sites. Geochemical data from interesting features between the landing sites and from the significant hard-to-reach areas will not be available.
Reconnaissance geology	Simple instrumentation on a rover can provide a wealth of information on the structure, stratigraphy, geomorphology, and physical properties of the lunar surface in the environs of each post-*Apollo* site and on extended traverses between them.
Instrument deployment	Can place science instrument packages in areas that cannot be reached by other methods.
Search	Provides a means for investigating the nooks and crannies of the lunar surface in the search for evidence of lunar life or other phenomena.